



Universidad
Carlos III de Madrid

DEPARTAMENTO DE INFORMÁTICA

PROYECTO FIN DE CARRERA

INGENIERÍA TÉCNICA EN INFORMÁTICA DE GESTIÓN

DISEÑO DE UN SISTEMA DE ANÁLISIS Y MODELO DE RIESGOS DE DETECCIÓN DE CONTENIDO PEDÓFILO EN REDES SOCIALES.

AUTOR: JULIA BEJARANO AGUADO

TUTOR: JOSÉ MARÍA ÁLVAREZ

LEGANÉS, OCTUBRE DE 2016

TÍTULO: DISEÑO DE UN SISTEMA DE ANÁLISIS Y MODELO DE
DETECCIÓN DE RIESGO DE CONTENIDO PEDÓFILO EN
REDES SOCIALES
AUTOR: JULIA BEJARANO AGUADO
TUTOR: DR. JOSÉ MARÍA ÁLVAREZ

EL TRIBUNAL

PRESIDENTE: VALENTÍN MORENO

VOCAL: JESÚS POZA

SECRETARIO: ROY MENDIETA

REALIZADO EL ACTO DE DEFENSA Y LECTURA DEL PROYECTO FIN DE CARRERA EL DÍA 25 DE OCTUBRE DE 2016 EN LEGANÉS, EN LA ESCUELA POLITÉCNICA SUPERIOR DE LA UNIVERSIDAD CARLOS III DE MADRID, ACUERDA OTORGARLE LA CALIFICACIÓN DE

VOCAL

SECRETARIO

PRESIDENTE

A mis padres

A mis hermanos

Y a mis hijos

Agradecimientos

A Jose María , que además de tutor del proyecto, has sido como el rayo de luz que ilumina un camino dominado por la oscuridad ; con su paciencia, por haber confiado en mí, por sus ánimos, por su optimismo has resultado ser el mejor apoyo en este largo y difícil camino para mí.

A mi amiga Patricia, por tu inestimable ayuda en la consecución de este proyecto. Gracias por haberme tendido una mano cuando tanto la necesitaba, por el cariño que siempre me has dado y porque ambas sabemos lo duro y difícil que resulta involucrarse en viajes de esta envergadura.

A mis padres, Fernando y Francisca, porque sé que siempre estarán cuando les necesite, y no hay palabras con las que expresar mi amor por todo lo que me han dado.

Y finalmente, quiero dar las gracias a la “niña” que todavía habita en mí, y que sigue mostrando el mismo afán de superación y curiosidad como cuando tenía 4 años y leía y escribía, o cuando jugaba al ajedrez con 5, o cuando con 8 años ayudaba a su padre con trabajos de Pedagogía de autores como Rousseau, Piaget, Maria Montessori, etc. Gracias por no haberme abandonado nunca, y por seguir de mi mano ahuyentando las disciplinas y obligaciones que impone el devenir de esta encorsetada vida, y seguir creyendo que la mente es una entidad cautivadora que no tiene edad, sexo ni fin.

“Siento una enorme gratitud por todos los que me dijeron “NO”. Gracias a ellos, lo hice yo mismo” A. Einstein, Físico alemán

*“Where there is a will, there is a way”,
Popular saying*

Resumen

El uso de las nuevas tecnologías ha experimentado un incremento considerable en los últimos años. Las tecnologías de la información y la comunicación (TICs) se han aplicado en distintos ámbitos con el objetivo de lograr un beneficio para los ciudadanos en cualquier esfera de sus vidas. Aspectos como que la red favorece una comunicación más desinhibida, entre personas que pueden hallarse en cualquier lugar y a cualquier hora, que ofrece la posibilidad de crear un “yo electrónico” que no tiene que parecerse al “yo real”, sin requerir coincidencia ni de lugar ni de tiempo son referidas como características de la red que pueden favorecer que la misma sea un lugar elegido por los depredadores sexuales para actuar. Así, junto a las ventajas que la red ofrece como espacio global de intercambio de información, existen peligros inherentes a la misma, donde encontramos como usos indebidos de las TICs, los llamados ciber-delitos que tienen como víctimas a los menores, y entre los que cabe destacar el ciberbullying, sexting y grooming. Es este último, el delito más usado por pedófilos en las redes sociales.

Los ciber-delitos cuyos tipos delictivos tienen como víctimas a menores de edad, tales como la actividad pedófila online (grooming), son de gran y creciente preocupación para la sociedad. Un análisis profundo de los actos de diálogo pedófilos en Internet ha implicado diseños de soluciones efectivas que discriminen de manera automática las conversaciones engañosas de las normales. Creemos que un estudio más profundo de la conversación pedófila puede revertir en sistemas de vigilancia más sofisticados y robustos que la mayoría de los sistemas actuales que se sostienen en un procesamiento

superficial del número de palabras o del reconocimiento de palabras clave: planteamos un estudio de las conversaciones pedófilas desde la perspectiva de "la teoría del

grooming online" y un desarrollo de una serie de análisis sobre varios actos de diálogo pedófilos con el objetivo de conseguir unos patrones de reconocimiento útiles.

La pedofilia se puede definir como la atracción sexual de un adulto hacia un niño. El DSM-IV-TR la incluye dentro del apartado de los trastornos sexuales. Las personas que presentan este trastorno muestran variabilidad en diferentes características, entre ellas la presencia de trastornos y rasgos patológicos de personalidad.

El objetivo del trabajo es revisar estudios publicados en la última década sobre la patología de la personalidad en pedofilia para hacer un análisis de detección de conversaciones con contenido pedófilo en las redes sociales y una propuesta de un Modelo de Riesgos. El trabajo evidencia que aunque existen estudios que muestran en pedófilos unos rasgos y unos trastornos de personalidad predominantes, es difícil presentar unos rasgos que sean característicos de este grupo, debido a que la mayoría de estudios son poco concluyentes, además del reducido volumen de trabajos y a la necesidad de controlar otras variables.

Palabras clave

TIC (Tecnologías de la Información); Pedofilia; Psicopatología; Ciberbullying; Sexting; Grooming; Pederastia versus Pedofilia; D SM-IV-TR (Manual diagnóstico y estadístico de los trastornos mentales).

Abstract

The use of new technologies has experienced an important rise in the last years. The purpose of promoting the development and the use of the Information and Communication Technologies (ICTs) is to provide benefits to the population in any areas of their lives. Issues such as the fact that the use of Internet promotes an easygoing communication between people who can be reached online anywhere and at any time the fact that Internet communication offers the possibility of creating “a virtual identity” which does not need to match the “real identity”, nor coincide in time nor place, are considered characteristics of the Net, often making it the chosen field of attack for sexual predators.

Therefore, together with the benefits offered by the Net as a global area of information exchange, there are inherent risks within it, where as examples of misuses of the ITCs, we find the so called Cyber-crimes targeting children, i.e. cyberbullying, sexting and grooming (online pedophile activity). The latter being the most common offence committed by pedophiles on the social networks.

Cyber-crime targeting children such as online pedophile activity are a major and a growing concern to the society. A deep understanding of predatory chat conversations on the Internet has implications in designing effective solutions to automatically identify malicious conversations from regular conversations. We believe that a deeper understanding of the pedophile conversation can result in more sophisticated and robust surveillance systems than majority of the current systems relying only on shallow processing such as simple word-counting or key-word spotting: we study pedophile

conversations from the perspective of online grooming theory and perform a series of analysis on several pedophile chat conversations to gain useful insights and patterns.

Pedophilia is defined as the sexual attraction of an adult towards a child; the DSM-IV-TR includes it in the section on sexual disorders. People with this disorder show variability in different traits, including the presence of disorders and pathological personality traits. The aim of this study is to review studies published in the last decade on the pathology of pedophilia personality in order to elaborate an analysis of detection of pedophile chat dialogues in social networks and a proposal of a Model of Risks. Since most studies are inconclusive, the low volume of investigations and the need to control other variables, the work highlights that although there are studies that show in pedophiles some predominant traits, and personality disorders, is difficult to present some features that are characteristic of this group

Keywords

Information and Communication Technologies (ICTs); Pedophilia; Psychopathology; Cyberbullying; Sexting; Grooming; Pederasty versus Pedophilia; DSM-IV-TR (Diagnostic and Statistical Manual of Mental Disorders).

Índice general

INTRODUCCIÓN	1
1.1. Contexto	2
1.2. Motivación	4
1.3. Objetivos	5
1.4. Estructura de la memoria	6
ESTADO DEL ARTE	8
2.1. Técnicas de Análisis de Actos de Diálogo Aplicadas en la Detección de Pedófilos	9
2.2. Agentes Software para la Detección de Conversaciones con Contenido Pedófilo	13
DISEÑO Y DESARROLLO	16
3.1. Conceptos Introductorios	17
3.2. Perfiles de la Víctima (Menor), Pedófilo y Conversación en las Redes Sociales.	18
3.2.1 Perfil Víctima	19
3.2.2 Perfil agresor sexual	21

3.2.3	Perfil Conversación	25
3.3.	Las TICs al Servicio de la Detección de Conversaciones de Riesgo Pedófilo: Propuesta de un Modelo de Riesgos	28
3.4.	Implementación en desarrollo con Técnicas SVMs	30
	VALIDACIÓN PRÁCTICA	31
4.1.	Validación Modelo de Riesgos	32
	MARCO LEGAL	37
5.1.	Las TIC y el Derecho Penal: Child Grooming y otras Formas de Abuso	38
5.2.	El delito de <i>Grooming</i> (art. 183 bis del Código Penal)	39
5.2.1.	El artículo 183 bis	39
5.2.2.	Requisitos típicos del delito de grooming	40
5.3.	Conclusiones	41
	PRESUPUESTO Y PLANIFICACIÓN	43
6.1	Planificación	44
6.2	Costes	46
6.2.1	Costes de material	46
6.2.2	Costes de recursos humanos	46
6.2.3	Costes indirectos	47
6.2.4	Presupuesto Total del Proyecto y Ganancia Estimada	48
	CONCLUSIONES Y LÍNEAS DE TRABAJO FUTURAS	49
7.1	Conclusiones	50
7.2	Líneas de trabajo futuras	50
	BIBLIOGRAFÍA	51

Índice de tablas

Tabla 1. Ejemplo de codificación de variables en el Modelo de Riesgos	32
Tabla 2. Costes de material	46
Tabla 3. Desglose costes de recursos humanos del proyecto	47
Tabla 4. Costes indirectos asociados al proyecto	47
Tabla 5. Presupuesto total y ganancia estimada del proyecto.....	48

Índice de figuras

Figura 1. Arquitectura de agentes inteligentes para el procesamiento de información en la detección de explotación de niños	9
Figura 2. Categorías de actos de Diálogos en conversaciones de chat	10
Figura 3. Grupos de edad vs longitud de frases	11
Figura 4. Grupos de edad vs Longitud de documento de chat	11
Figura 5. Componentes del sistema ChatTrack: Chatprofile y ChatRetrieve	12
Figura 6. Diseño detallado del sistema	13
Figura 7. Estructura de las áreas de las que hace uso Negobot	15
Figura 8. Variables que definen el perfil de pedófilo en el Modelo de Riesgos	21
Figura 9. Los seis estado del Child-Grooming	22
Figura 10. Símbolos usados por pedófilos	23

Figura 11. Variables que definen el perfil de pedófilo en el Modelo de Riesgos25

Figura 12. Variables que definen el perfil de conversación en el Modelo de Riesgos27

Figura 13. Modelo de Riesgos28

Figura 14. Conjunto de variables del Modelo de Riesgos29

Figura 15. Matriz de variables sin riesgo33

Figura 16. Matriz de variables con riesgo máximo34

Figura 17. Matriz de variables según un ejemplo práctico.....35

Figura 18. Diagrama de Gantt45

Capítulo 1

Introducción

El primer capítulo de este Proyecto Fin de Carrera está destinado a proporcionar una visión global del mismo. Comenzaremos por presentar el contexto donde queda encuadrado nuestro estudio, los diferentes aspectos que han motivado el desarrollo y los objetivos que se pretenden alcanzar en él. Finalmente se describe como han sido estructurados los diferentes contenidos del presente documento.

1.1. Contexto

Internet es un medio con enormes posibilidades para encontrar información, aprender, expresar opiniones y comunicarse, así como un canal de ocio. Pero Internet también puede ser un lugar donde existen ciertos **riesgos y peligros** que se acentúan en los menores, al ser éstos mucho más vulnerables que los adultos. Por tanto hay que conocer los riesgos y peligros y poner los medios necesarios para proteger de ellos a los menores. Pero, ¿a qué riesgos se pueden enfrentar los menores? Los posibles riesgos que un menor puede encontrar en Internet se centran en tres tipos:

- **Contenidos inapropiados.** Al navegar por Internet pueden encontrarse con contenidos no adecuados para su edad, como por ejemplo páginas con un lenguaje inadecuado, con contenidos para adultos o contenidos violentos, etc.
- **Privacidad.** Todo el mundo debe tener cuidado a la hora de facilitar datos privados en Internet. Este riesgo se ve incrementado en el caso de los menores ante su mayor ingenuidad. Deben de tener claro que no deben facilitar información personal que pueda poner en riesgo a sus familiares, compañeros y a ellos mismos.
- **Discriminación, abusos y acosos.** Esto sucede, sobre todo, a través de programas de mensajería, chats, foros, etc.

Este último contenido que mencionamos constituye el núcleo central de nuestro estudio, pues la lista de peligros y riesgos que acechan a los menores en la Red no resulta nada corta. Además del acoso sexual, el ciberbullying y los depredadores también llamado pedófilos, constituyen hoy en día el problema más preocupante, pues motivos suficientes hay para preguntarse cuán preparado está un niño o prepuberto para moverse por las redes sociales con cierta seguridad cuestionándose cuál es la edad más recomendable para que un niño, ya ni siquiera hablamos de adolescentes, comience a usarlas, preocupaciones que traen de cabeza a muchos padres, arrastrados por un entorno que ya es digital desde la más tierna infancia de sus hijos.

La terminología “pedofilia” fue creada por primera vez por el psiquiatra Richard Von Krafft-Ebin en su libro de psicopatías sexuales [1]. Desde entonces ha existido un debate respecto a la trascendencia de los términos pedofilia y pederastia, ambos términos reflejan distintos alcances de la atracción sexual hacia los menores de edad, mientras la pedofilia consiste en la atracción erótica por los niños, la pederastia es en sí misma el abuso sexual de ellos. Sin embargo, hoy en día el término pedofilia ha perdido su carácter inocuo, y se refiere directamente a la práctica abusiva y delictiva, de modo ambos términos, pedofilia y pederastia son términos comunes y a veces, incluso intercambiables.

La pedofilia ha estado presente desde tiempos arcaicos siendo considerada una práctica normal la unión entre adultos y menores de edad, a través de acuerdos matrimoniales a los que llegaban las familias de las partes involucradas. Pero con el transcurrir del tiempo estos acuerdos entre familias se designaron ilegales debido a que las leyes actuales impiden el matrimonio con menores de edad, además de ser considerado como un delito grave ya que hoy en día la pedofilia conlleva a un abuso o a una práctica inmoral que afecta física y psicológicamente a los menores de edad que por lo general son abusados sexualmente sin consentimiento propio.

En la actualidad, pederastia y pedofilia son los dos términos que se relacionan con los agresores implicados en acciones o delitos sexuales con menores. Aunque, desde el punto de vista clínico, el DSM-IV-TR sólo utiliza el término de pedofilia. Esta patología se incluye dentro del apartado de los trastornos sexuales, compartiendo capítulo con el resto de las parafilias. La gama de actos sexuales que los pedófilos pueden realizar con los niños es amplia. Puede ir desde actividades como el exhibicionismo, o el voyerismo, a otras como realizar caricias, frotar sus genitales contra un niño, masturbación en presencia de estos, sexo oral y penetración anal o vaginal [2] [3] Aunque generalmente estas personas no usan la fuerza para implicar a los niños en actividades sexuales sino que usan varias formas de manipulación psíquica y desensibilización (progresión de tocamientos de inocuos a inapropiados, mostrando pornografía a los niños), justificando sus comportamientos como actos con valor educativo para los niños o actos de los que el niño obtendría placer [2]; [4].

Hay que tener en cuenta que no todas las personas con pedofilia son iguales en sus preferencias, edad de inicio, ni actúan con las mismas estrategias. Mientras que unos son violentos, otros son seductores de niños. Unos prefieren a los niños y otros prefieren a niñas, unos a niños de menos edad y otros a niños algo mayores [5]. Aunque su curso es habitualmente crónico a partir de su inicio [6] y la atracción sexual hacia los niños empieza en la pubertad y adolescencia, muchos pedófilos también la pueden desarrollar a edades adultas [2]. Estos datos se pueden tomar como una muestra de la gran variabilidad de características que existe entre personas que comparten un diagnóstico de pedofilia. Debido al gran interés social que suscita este tema, se han realizado estudios que han intentado hacer una aproximación a diferentes características psicológicas que pudieran estar relacionadas con esta patología.

En cuanto a variables de personalidad, diferentes trabajos también recogen la dificultad de encontrar unos rasgos característicos o comunes de este tipo de patología. Así, varios autores han encontrado evidencia de trastornos y rasgos psicopatológicos en esta población. Algunos de éstos son el trastorno psicopático en hombres pedófilos. Rasgos de personalidad con dominancia de signos de conducta pasiva-dependiente [7]. Trastornos de personalidad narcisista, visión distorsionada de los otros y necesidades primitivas de dependencia en pedófilos encarcelados [8]. Mientras que otros no encuentran suficiente evidencia de una psicopatología específica y claramente definible, por lo que definen a los pedófilos como personas generalmente “normales” pero muy hábiles para planear su conducta delictiva y negar su existencia [9] y que muestran una inhabilidad para ver su actividad sexual como necesaria de tratamiento, además de una resistencia a cambiar esta conducta [10].

Teniendo en cuenta la relevancia del tema de la pedofilia, agresión y abuso a menores y la dificultad de encontrar rasgos de personalidad característicos, se pretende realizar una herramienta de detección y aproximación a la patología de la personalidad en pedofilia en redes sociales, ya que dentro del estudio de la personalidad sea cual sea la muestra, las investigaciones más numerosas son las referentes a rasgos patológicos de la personalidad y a los trastornos de personalidad comórbidos.

1.2. Motivación

En la actualidad las redes sociales han llegado a ser de gran importancia para la sociedad; es el caso de *Twitter* que actualmente cuenta con más de 500 millones de personas conectadas en tiempo real alrededor del mundo: Es gracias a esta red que más de un *microblogging* se ha convertido en una fuente de noticias al último minuto. Los empresarios, famosos, investigadores, instituciones gubernamentales y privadas, además de un sinnúmero de usuarios que también se conectan a esta red para compartir sus experiencias e inquietudes o noticias de última hora. Toda esta gran cantidad de información que se encuentra en la red, así como su rápido crecimiento, hacen que muchos quieran indagar un poco más sobre algunos perfiles que pudieran ser importantes para la investigación moderna y también ser parte de nuevos descubrimientos, soluciones o simplemente por encontrar su potencial dentro de la red.

Por todos estos motivos, la pedofilia se nos presenta como una actividad que se ha incrementado con el desarrollo de Internet. Una de las razones es el anonimato y la facilidad de hacer contactos, tanto con menores de edad como con otros pedófilos. Varios proyectos e iniciativas para investigar este fenómeno, se han desarrollado a nivel mundial. Las agencias gubernamentales encargadas de minimizar este fenómeno utilizan primordialmente dos estrategias: páginas ficticias que hacen el papel de señuelo para detectar usuarios pedófilos, y agentes de policía que haciéndose pasar por niños logran engañarlos para su posterior arresto.

Un estudio realizado en los Estados Unidos puso de manifiesto que uno de cada cinco niños y niñas que se incorporan al mundo de las redes sociales reciben solicitudes por parte de extraños para entablar relaciones sexuales. En otro estudio se reveló que cada año hay unos 104.000 niños y niñas víctimas de abuso sexual. Un comunicado de prensa de la UNICEF, menciona que en América Latina, se estima que cada año más de 2 millones de niños, niñas y adolescentes son abusados sexualmente. Los niños corren mayor riesgo de sufrir violencia física que las niñas, aunque éstas corren un mayor riesgo de sufrir violencia sexual, abandono y prostitución forzada. De todos estos datos, y gracias a las labores de seguimiento de innumerables organismos, no son suficientes para poner freno a estos delitos. Según K. F. Durkin, en "Misuse of the Internet Pedophiles" son cuatro las actividades que los pedófilos desarrollan en Internet [11]:

- El tráfico de pornografía infantil e historias en textos que manejen un contenido pedofílico. Esto lo hacen por medio de correos electrónicos, páginas web, chat rooms.
- Localizar niños vía chat para acosarlos. Esto lo hacen en los chat rooms en donde inician una relación con los niños o con las personas que cuidan a estos niños.
- Enganchar en conversaciones inapropiadas de tipo sexual con los niños donde hay intercambio de archivos.
- Comunicación con otros pedófilos en quienes encuentran apoyo. Ellos necesitan la aprobación de sus colegas para el intercambio de experiencias y dificultades en el manejo de sus sentimientos ya que socialmente son rechazados. Este tipo de comunicaciones generalmente involucra tecnologías como: newsgroups, páginas Web, email, mensajería instantánea y chat rooms.

Varias iniciativas se han desarrollado para la protección de los niños y sus padres. En cuanto al peligro de las redes sociales, y a pesar de ellas, se ha detectado la existencia, en primer lugar, de un problema de “educación en cuanto al peligro de Internet”, ya que muchos niños no han sido prevenidos de los peligros y riesgos que entrañan las redes sociales por no haber recibido unas pautas y consejos sobre la seguridad de Internet por parte de sus padres; en segundo lugar, las iniciativas desarrolladas para la protección de los niños en cuanto a sistemas o herramientas capaces de la detección de niños explotados, a través de actos de diálogos en un chat room, no son los suficientemente eficaces y precisos como para ser capaces de una detección y bloqueo precoz de conversaciones que identifiquen a la víctima y al presunto pedófilo.

Por todo esto, el objetivo de este Proyecto Fin de Carrera es precisamente desarrollar un modelo, así como una herramienta de detección y aproximación a la patología de la personalidad en pedofilia en redes sociales que permita procesar información procedente de actos de diálogo en chat rooms con la finalidad de detección de “posibles” conversaciones entre un menor y un pedófilo.

1.3. Objetivos

Las Redes Sociales son básicamente unidades virtuales, donde existen usuarios que interactúan unos con otros a través de Internet. Se agregan perfiles que tienen algo en común y la plataforma que se utiliza permite conectar gente que se conoce realmente o que desea conocerse por algún atributo que encuentren en común, según sea el caso de la plataforma: en algunas se pueden compartir fotos, vídeos, centralizar recursos, foros de discusión, mensajes, etc. Existen actualmente muchos estudios sobre las redes sociales, que se basan en gran medida en estudios individuales sobre el comportamiento de las personas. Estos estudios también están enfocados en mecanismos que utiliza la red para unir todo este ambiente, donde analizan y describen los niveles de la red social.

El objetivo principal de este Proyecto Fin de Carrera es el diseño de una herramienta de aproximación y detección a la patología de la personalidad en pedofilia en redes sociales que permita procesar información procedente de actos de diálogo en chat rooms con la finalidad de detección de conversaciones entre un menor y un “presunto” pedófilo.

Este objetivo global se operacionaliza en los siguientes objetivos más específicos:

- Planificación y diseño de las etapas en las que se dividirá el presente Proyecto Fin de Carrera.
- Estudio de los estudios ya existentes relacionados con nuestro tema sobre la Pedofilia con el fin de obtener una panorámica de la situación actual en la que nos encontramos relativa a la detección de actos de diálogo pedófilos en redes sociales.
- Desarrollo de un módulo software que permita detectar conversaciones entre un menor y un “presunto agresor sexual” a partir de la información recogida en archivos de las redes sociales.
- Validación del módulo desarrollado de forma individual y posteriormente de forma global en el ámbito de la aplicación Web.
- Documentación del desarrollo realizado.

1.4. Estructura de la memoria

La organización de los contenidos del presente Proyecto Fin de Carrera se detalla a continuación:

- **Capítulo 1: Introducción**

El primer capítulo de este Proyecto Fin de Carrera está dedicado a proporcionar una visión global del mismo. En este capítulo se presenta el contexto donde se enmarca nuestro estudio, los diferentes aspectos que han motivado el desarrollo de este PFC, los objetivos que se pretenden alcanzar en él y la estructura del presente documento.

- **Capítulo 2: Estado del Arte**

El propósito de este capítulo es obtener una visión global de todas aquellas tecnologías y herramientas relacionadas con el presente Proyecto Fin de Carrera. Así, se revisan las tecnologías basadas en el uso de técnicas de minería de texto, utilizando una serie de algoritmos que ayudan a determinar ciertos patrones del comportamiento humano, arquitecturas de agentes inteligentes para el procesamiento de información en la detección de explotación de niños

- **Capítulo 3: Diseño y Desarrollo**

En este capítulo se describe el proceso de diseño y su desarrollo e implementación para la consecución de un Modelo de Riesgos. Además, se proporciona una visión global de la aplicación con el fin de contextualizar cada una de las variables de los perfiles explicados detalladamente en el dicho capítulo.

Asimismo, el objetivo de este capítulo es presentar los aspectos más conceptuales y técnicos del desarrollo. Se detalla la implementación de cada uno de los perfiles en el orden en que fueron desarrollados para permitir comprender el proceso de forma clara y sencilla.

- **Capítulo 4: Validación**

Este capítulo detalla la comprobación realizada con un ejemplo práctico de un vector de variables de los perfiles, junto con la generación de la tabla Excel generada del Modelo de Riesgos, para comprobar el correcto funcionamiento del módulo software desarrollado.

- **Capítulo 5: Marco legal del delito de Child Grooming.**

En este capítulo se ofrece una breve referencia a la legislación existente sobre pedofilia, para facilitar una aproximación al marco normativo español dirigido a regular este tipo de delitos.

- **Capítulo 6: Presupuesto y Planificación**

Se presenta la planificación y fases de desarrollo del Proyecto Fin de Carrera así como el presupuesto del mismo, desglosándolo en los correspondientes costes de material y costes de recursos humanos.

- **Capítulo 7: Conclusiones y Líneas de Trabajo Futuras**

En este capítulo final, se exponen las conclusiones obtenidas tras la ejecución del presente Proyecto Fin de Carrera, así como posibles trabajos futuros.

Capítulo 2

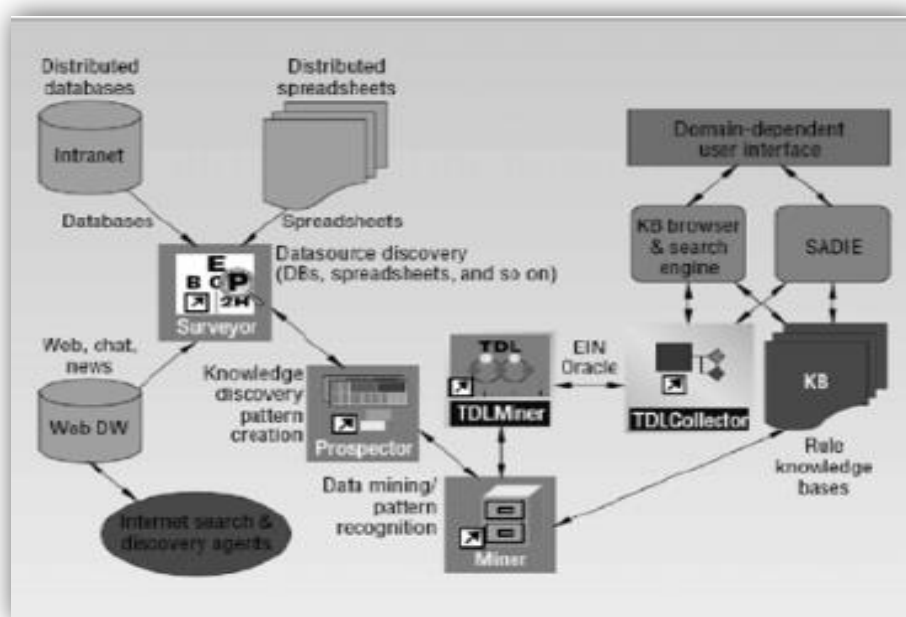
Estado del arte

El propósito de este capítulo es obtener una visión global de todas aquellas tecnologías y herramientas relacionadas con el presente Proyecto Fin de Carrera. Varias iniciativas se han desarrollado para la protección de los niños y sus padres, como online Safety Project. Sin embargo, a pesar de la existencia de ésta y otras, se ha detectado que muchos niños no han sido educados en cuanto al peligro en Internet. Por ejemplo, se ha reportado que muchos niños no han recibido ningún consejo para la seguridad de Internet de parte de sus padres y por lo tanto están desprevenidos acerca de la identidad de algunos miembros en las salas de chat. En primer lugar, se hará una Revisión sobre las técnicas de análisis de conversaciones aplicadas en la detección de pedófilos en Internet. Dentro de las técnicas utilizadas para identificar a la víctima y al presunto pedófilo se encuentran las Técnicas en el procesado y Análisis Léxico de las colecciones de conversaciones escogidas. En segundo lugar, haremos una recapitulación sobre los diversos mecanismos existentes en las redes sociales, desarrollados para la identificación de agresores sexuales.

2.1. Técnicas de Análisis de Actos de Diálogo Aplicadas en la Detección de Pedófilos

Steve G. Romaniuk [12], propone un sistema basado en agentes inteligentes como apoyo a las oficinas de inteligencia, para detección de niños explotados, el sistema inteligente se basa en el descubrimiento del conocimiento y predicción de información para fuentes de datos en una intranet de y de análisis web, búsqueda y descubrimiento en Internet, según el sistema de agentes desarrollados mostrados en la siguiente figura (Figura 1):

Figura 1. Arquitectura de agentes inteligentes para el procesamiento de información en la detección de explotación de niños



Fuente: Beltran, A. (2010). Identificación de conversaciones de pederastas en Internet (p. 2).

En sus estudios, propone una arquitectura de un sistema de indexación incremental para guardar y recuperar información de los chats room y de mensajería instantánea que soporta diferentes formas de consultas.

Recientemente (2007), E. N. Forsythand y C. H. Martell [13] desarrollaron un corpus etiquetado con información sobre léxico, sintaxis y discurso basado en conversaciones en línea de chat. Este corpus que consta de más de 475.000 post fue automáticamente etiquetado a través de redes neuronales en 15 categorías de actos de diálogo (Figura 2).

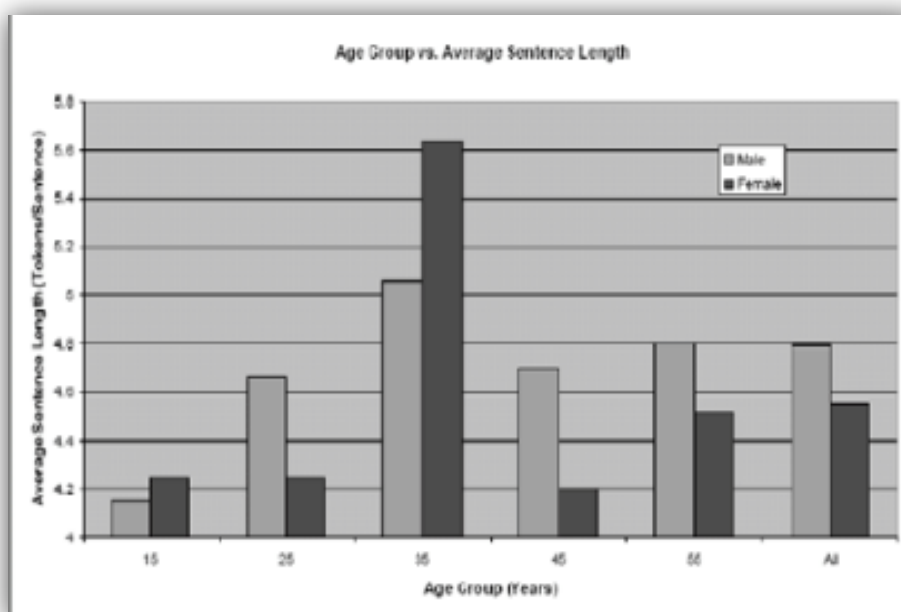
Figura 2. Categorías de actos de Diálogos en conversaciones de chat

Classification	Example
Accept	yeah it does, they all do
Bye	night ya'all.
Clarify	i meant to write the word may.....
Continuer	and thought I'd share
Emotion	lol
Emphasis	Ok I'm gonna put it up ONE MORE TIME 10-19-30sUser37
Greet	hiya 10-19-40sUser43 hug
No Answer	no I had a roommate who did though
Other	0
Reject	u r not on meds
Statement	Yay...democrats have taken the house!
System	JOIN
Wh-Question	11-08-20sUser70 why do you feel that way?
Yes Answer	why yes I do 10-19-40sUser24, lol
Yes/No Question	cant we all just get along

Fuente: Beltran, A. (2010). Identificación de conversaciones de pederastas en Internet (p. 2).

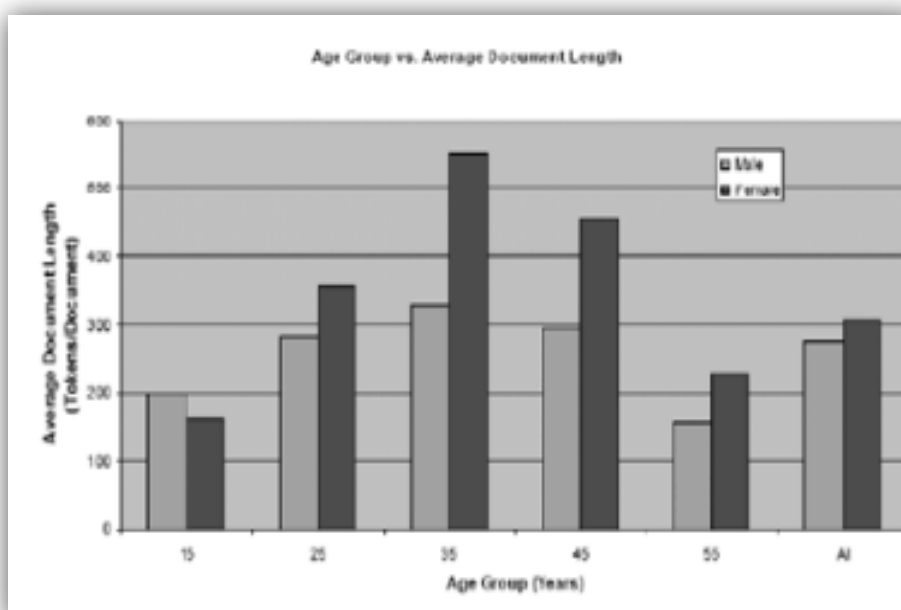
En otro trabajo sobre el perfilamiento de usuarios en Internet, Jin Lin aborda el problema para la identificación de género y edad de las personas en los chat basado en los logs [14]. En el estudio realizado se encontró que los hombres tienen un mayor promedio en la longitud de las frases que las mujeres, quienes tienden a escribir frases más cortas pero a crear documentos de chat más extensos (Figura 3 y 4).

Figura 3. Grupos de edad vs longitud de frases



Fuente: Beltran, A. (2010). Identificación de conversaciones de pederastas en Internet (p. 2).

Figura 4. Grupos de edad vs Longitud de documento de chat



Fuente: Beltran, A. (2010). Identificación de conversaciones de pederastas en Internet (p. 3).

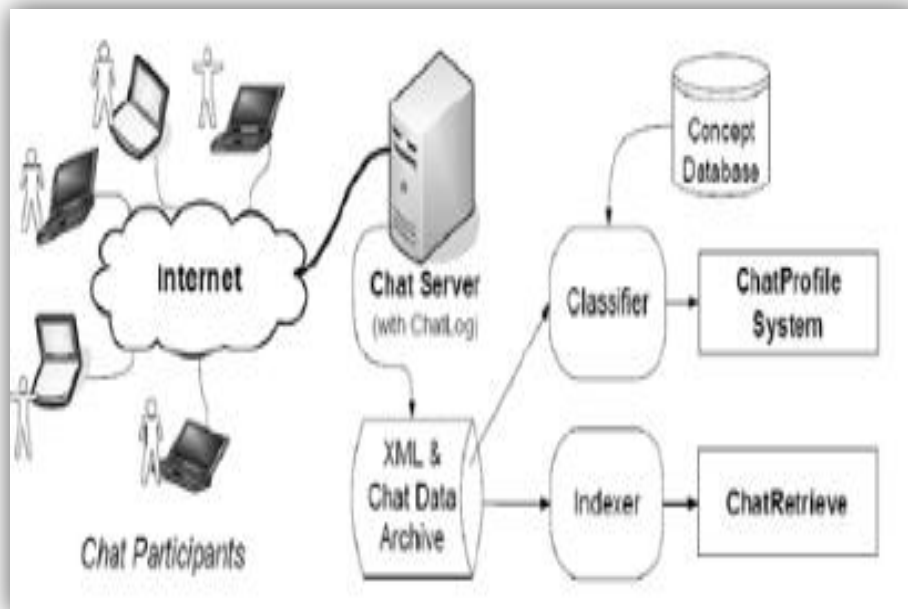
En otro trabajo realizado por **N. Pendar** [15] se presentan los resultados de un estudio que utiliza las técnicas de categorización automática de conversaciones escritas de un chat en la identificación en redes sociales de pedófilos. Utiliza un corpus disponible en www.perverted-justice.com sobre conversaciones pedófilas y a través de técnicas de aprendizaje automático utilizando SVM y clasificadores k-NN, por lo que dado un archivo particular que contenga las

conversaciones de un pedófilo y su víctima, es posible distinguir automáticamente los dos lados de la conversación con una alta precisión.

Además de estos sistemas de procesamiento y análisis léxico, se dispone de otras herramientas de muestreo de conversaciones de un chat room en el que se maneje un posible ambiente pedófilo como:

- **Butterfly** [16] es un agente de software desarrollado para muestrear los grupos de conversación disponibles en un chat room. Este software se basa en el IRC (Internet Relay Chat) donde este agente se introduce en los chat por un período de 30 segundos durante los cuales puede identificar el tipo de lenguaje que los usuarios están empleando.
- **ChatTrack** [17] es una herramienta creada para la detección de una posible actividad pedofílica con la identificación de menores, lenguaje de adultos y la detección de adultos haciéndose pasar por niños (Figura 5).

Figura 5. Componentes del sistema ChatTrack: Chatprofile y ChatRetrieve



Fuente: Beltran, A. (2010). Identificación de conversaciones de pederastas en Internet (p. 3).

- **PieSpy** [18] es un robot espía IRC que intenta determinar las relaciones en un conjunto de canales IRC de las salas de chat y los muestrea visualmente. La heurística que utiliza para determinar una relación entre dos usuarios incluyen la de detectar, dado un Nick-name, con quienes éste ha establecido conversación y la rapidez de los mensajes y las respuestas.

Con todas estas investigaciones y estudios se pretende que con la consideración de actos de diálogo, los clasificadores de conversaciones de pedófilos en Internet tengan un alto nivel de precisión y efectividad.

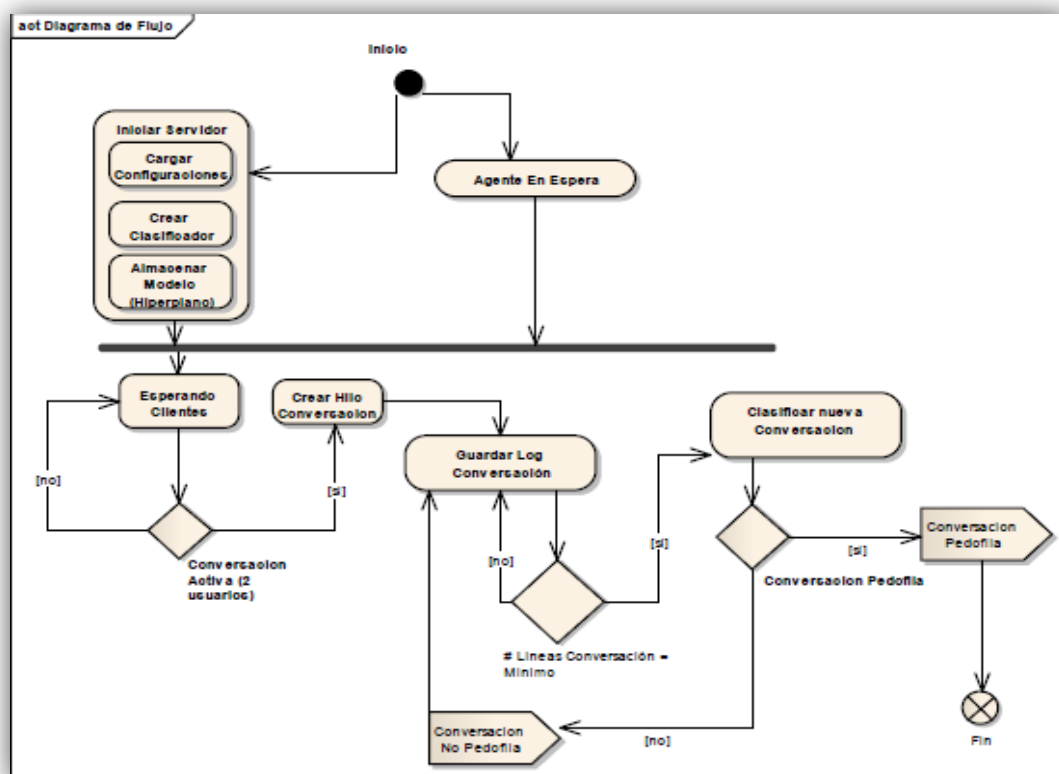
2.2. Agentes Software para la Detección de Conversaciones con Contenido Pedófilo

En un contexto donde, por un lado, la evolución de la tecnología y comportamientos sociales y, por otro, el crecimiento en las redes sociales de riesgos asociados a la infancia, fenómenos como el acoso en todas sus vertientes (ciberbullying, grooming, etc.) o la pornografía infantil, se han propuesto implementaciones de agentes software para identificar las conversaciones con contenido pedófilo que puedan presentarse en un chat, fundamentados en técnicas de procesamiento de lenguaje natural y algoritmos de aprendizaje automático supervisado. Entre ellos tenemos:

- **DISEÑO SISTEMA INTELIGENTE [9]**

Se propone una arquitectura para el sistema inteligente y su posterior implementación en la plataforma JADE (Java Agent Development Framework). Un esquema general del sistema es el que se muestra en la siguiente figura (Figura 6).

Figura 6. Diseño detallado del sistema



Fuente: Beltrán, A, y Ordóñez, S. (2014). Sistema inteligente para la detección de diálogos con posibles contenidos pedofílicos. Revista Virtual Universidad Católica del Norte, 42, 164-181 (p. 176)

El sistema inicial [12] colocando el servidor chat en línea, dicho servidor lee los archivos de configuración y de entrenamiento, crea el clasificador y queda a la espera de un cliente.

Inmediatamente un cliente se conecta al servidor, se crea un hilo para dicha conversación, este empieza a leer y guardar las líneas de la conversación en un archivo, apenas este archivo supere el número mínimo de líneas especificadas en el archivo de configuración, es enviado al clasificador para determinar si es una conversación con contenido pedófilo.

- **SWEETIE: LA NIÑA VIRTUAL**

Se trata de un programa virtual que emula a una niña filipina de diez años. La ONG sueca “Terre des Hommes” creó la representación digital denominada “Sweetie”, una niña virtual que permite identificar a los pedófilos. A través de la cámara web, “Sweetie” aparece en las salas de chat rooms. En el comienzo del proyecto, sólo en diez semanas, 20.000 hombres de todo el mundo, principalmente de Europa, Australia y Estados Unidos, se pusieron en contacto con la niña. Del total, 1.000 hombres fueron descubiertos y sus identidades entregadas a la Interpol, quienes encontraron las direcciones, números de teléfono y fotos a través de redes sociales y páginas web de los abusadores infantiles.

- **EL “NEGOBOT”: AGENTE CONVERSACIONAL BASADO EN TEORÍA DE JUEGOS**

El sistema “Negobot” busca identificar conductas agraviantes contra el menor. Es un agente conversacional que se hace pasar por un niño dentro de chats, redes sociales y otros servicios similares. Al ser un agente conversacional, hace uso de técnicas de procesamiento de lenguaje natural (PLN) y recuperación de la información (RI), así como de la Inteligencia Artificial y el Aprendizaje Automático. Sin embargo, la propuesta más innovadora de Negobot es la de considerar la conversación en sí misma como un juego, haciendo uso de la Teoría de Juegos, área de la matemática aplicada que utiliza modelos para estudiar interacciones en estructuras formalizadas de incentivos y llevar a cabo procesos de decisiones. Este robot utiliza un software de inteligencia artificial para chatear de manera realista e imitar el lenguaje que utilizan los adolescentes en las conversaciones; al principio es neutral pero adoptará cualquiera de las siete personalidades que tiene programadas dependiendo de la intensidad de las interacciones; también toma decisiones avanzadas para simular chats convincentes a medida que se desarrollan. Por lo tanto, el “Negobot” tiene la capacidad de tomar la iniciativa en conversaciones, así como recordar actos de diálogo y usuarios. Creado por investigadores españoles [19] donde la mayor contribución de este bot conversacional basada en Teoría de Juegos, es la de ofrecer una visión de una conversación como un juego, un juego con adversario. A raíz de esta visión, se ha construido un prototipo funcional que realiza conversaciones evaluando en cada momento las frases que este adversario está realizando con el fin de detectar pedófilos.

Figura 7. Estructura de las áreas de las que hace uso Negobot



Fuente: Laorden, C et al. DeustoTech Computing (2016). Negobot: Agente conversacional basado en teoría de juegos para la detección de conductas pedófilas.

Capítulo 3

Diseño y Desarrollo

3.1. Conceptos Introductorios

La primera pregunta que podríamos hacernos es **¿existen patrones de conducta en un pedófilo?** Las respuestas de prestigiosos psicólogos es que existen rasgos que se repiten en la personalidad de estas personas. Por eso ante dicha pregunta tendríamos que hacer un análisis detallado de *la Psicología del Pedófilo*. Para ello vamos a estructurar nuestro análisis de la siguiente forma:

- **CONCEPTO PEDOFILIA**

La pedofilia es, en esencia, un *autoerotismo*, y lo que realmente le atrae no es un cuerpo infantil sino más bien lo que la niñez “simboliza”, es algo psicológico. El pedófilo utiliza el cuerpo del niño para masturbarse en él. De ahí el éxito de las redes sociales entre los pedófilos pues ofrece sin un cuerpo el sexo anónimo. Los niños en el ciberespacio son meras representaciones, a menudo nada más que foto o videos eróticos. Los varones agresores están socializados para dominar: los niños les resultan atractivos debido a su baja capacidad de dominación- relación con inmadurez, baja autoestima y agresividad del agresor. Si bien no existe un perfil acabado del abusador sexual, distinguiremos dos categorías:

- **Abusador situacional (PEDERASTA):** Es el denominado delincuente sexual. Este tipo de abusador puede ser diagnosticado como portador de trastornos de personalidad del tipo psicopático/antisocial, narcisista, esquizoide y sádico. Es impulsivo y su acto criminal es espontáneo. Su modus operandi suele ser la coerción, la manipulación y el uso de la fuerza.
- **Abusador preferencial (PEDÓFILO):** Los abusadores de niños/as preferenciales tienen una preferencia sexual definida por los menores. Su conducta sexual es compulsiva, cometen su “delito” por una necesidad que les lleva a desplegar actos ritualísticos sexuales; es mentiroso y manipulador, planea eficientemente su delito y evalúa sus experiencias; suele usar la tecnología moderna, como las redes sociales, paraíso para los pedófilos, desde donde le es sencillo engañar con falsos perfiles y obtener información útil para entender y manejar a sus víctimas. Aunque este tipo de desviación esté incluida en el DSM-IV como Trastorno de la Sexualidad, este tipo de abusador es crónico, y por lo tanto, es menos probable que pueda modificar su conducta desviada.

- **INDICADORES TEÓRICOS DE LA PSICOPATOLOGIA DEL PEDÓFILO**

La pedofilia es un trastorno sexual que está clasificado dentro de las “parafilias” [2] que vendrían a ser desviaciones sexuales donde la fuente de placer está caracterizada por fantasías, impulsos excitatorios o comportamientos que abarcan desde objetos no humanos hasta el sufrimiento o humillación de uno mismo, la pareja, niños o adultos no consensuales. La pedofilia es pues una parafilia en la que hay una atracción sexual intensa, urgente, recurrente, por los niños.

Muchos son los psicólogos y médicos que han estudiado el perfil del pedófilo, llegando a conclusiones como que sí existen rasgos que se repiten en la personalidad de estas personas.

Entre ellos podemos destacar *la introversión*, que hace referencia a la baja capacidad para establecer vínculo con sus pares y las relaciones que establecen se vinculan a personas más vulnerables como son los menores de edad. Son *psicológicamente inestables, baja autoestima, ansiedad y miedo al abandono*. La única forma de satisfacer su sexualidad es a través de un niño, de lo que se deduce que este rasgo de *personalidad reprimida*, es otro patrón de conducta que dirigen a la relación con niños y a una postura de permanente conquista hacia ellos. El estudio del perfil psicológico del pedófilo nos acerca al concepto de *distorsiones cognitivas*, que son un sistema de creencias disfuncionales, como por ejemplo, las caricias sexuales no son en realidad sexo, y por ello, no se hace ningún mal; la falta de resistencias físicas significa que el niño desea contacto sexual; los niños no lo dicen debido a que les gusta el sexo. Un innumerable número de excusas que minimizan o niegan su responsabilidad.

• ETIOLOGÍA DEL PEDÓFILO

La manifestación del trastorno del pedófilo suele comenzar a edad temprana (con fantasías y deseos hacia menores de edad), normalmente suelen haber sido niños abusados y con traumas en su infancia y es difícil un tratamiento psicológico-conductual con éxito, pues como hemos perfilado anteriormente tiene en su origen un problema de fantasía sexual que les domina convirtiéndose en una adicción.

Tras este breve análisis, nuestra pregunta es si existen ciertos patrones de conducta que quedan más o menos respondidos; ahora tenemos que aplicar dicho perfil a las redes sociales, y al uso que el pedófilo hace en contenido y forma.

3.2. Perfiles de la Víctima (Menor), Pedófilo y Conversación en las Redes Sociales.

Para nuestro Modelo de Riesgos vamos a elegir tres perfiles que podrían ser considerados como variables independientes y predictivas en la posible detección de conversaciones con contenido pedófilo, y que a su vez contendrían sus consiguientes sub-variables. Serían:

- **Perfil víctima:**
 - Situación familiar
 - 1) Conflicto familiar
 - 2) Domicilio
 - Información sensible del menor
 - 1) Edad
 - 2) Género
 - 3) Nick
 - 4) Nº Amistades(A)
 - 5) Nº Publicaciones/Fotos(NP)
 - 6) Registro Actividad(RA)
- **Perfil pedófilo**
 - Identificación
 - 1) Significado Nick
 - 2) Verificación Nick
 - 3) Identificación Foto

- 4) Símbolo
 - Tiempo Conexión
 - N° Peticiones Conexión
- **Perfil Conversación**
 - Características físicas de la Conversación
 - 1) Densidad conversacional
 - 2) Duración
 - 3) Horario
 - Análisis de contenido
 - 1) Léxico: Emoticonos, Abreviaturas, Lenguaje pre-adolescente
 - 2) Semántico: Términos con contenido de carga sexual
 - 3) Sintáctico: Corpus en Base de Datos www.perverted-justice.com.
 - Etapas del proceso ON-LINE GROOMING
 - 1) Etapa 1: Acercamiento al menor
 - 2) Etapa 2: Toma de confianza
 - 3) Etapa 3: Sexual

A continuación desarrollamos las características que definen cada uno de los perfiles según se han incorporado al Modelo de Riesgos.

3.2.1 Perfil Víctima

La premisa de la que partimos, en nuestro diseño y desarrollo sobre patrones de conducta pedófilas en las redes sociales, es que normalmente *“el abuso sexual no es fruto de la casualidad”*.

En estos días, las plataformas on-line, específicamente las redes sociales, se convirtieron en el sitio perfecto para los pedófilos. Redes sociales como Facebook, Twitter, Google+ y Badoo esconden, con su anonimato, a pedófilos que buscan generar amistad con menores de edad e incluso establecer vínculos emocionales. Obtener “Me gusta” en redes sociales es parte vital de la vida de niños y jóvenes. Según un estudio realizado por Intel Security, la principal conclusión a la que se llegó es que más del 57% de niños consideran importante para ellos el número de “me gusta” obtenido por sus publicaciones, de entre las que destacan fotos personales, seguidas de selfies y fotos de grupos. El estudio también entregó datos dicientes. Para los preadolescentes lo peor que puede sucederles es que otras personas puedan acceder a información personal privada sin su consentimiento. Con mucha frecuencia, el abuso sexual del menor en redes sociales por parte de pedófilos se da en niños que han sufrido o sufren adversidades dentro de un rango que va desde niños que sufren relaciones conflictivas con sus padres inmersos en un conflicto familiar, hasta niños que padecen privaciones, desde indiferencias, maltrato emocional, e incluso abuso sexual por parte de sus padres. Actualmente, muchos estudios han demostrado que los padres que abusan de sus hijos suelen ingerir sustancias tóxicas, o padecen problemas psiquiátricos, tienen peleas maritales, son padres muy punitivos y distantes. Por sí solas, estas condiciones originan daños psicológicos, pero también es ésta inestabilidad emocional la que sitúa al niño en situación de riesgo porque son escasamente atendidos a la vez que están condicionados a aceptar la violencia y la victimización, y es por ello se vuelven vulnerables a las estrategias de los pedófilos en las redes sociales, quienes saben que tanto *las carencias afectivas como la inestabilidad emocional* del

menor son un caldo de cultivo propicio para lograr su objetivo. Todo esto se referirá en nuestro Modelo de Riesgos como la variable Situación Familiar.

Nuestro estudio se va a centrar en la detección de menores con edades menores de 13 años. Por otro lado, estudios recientes han alertado que los niños son más vulnerables al abuso sexual en redes sociales, y las niñas a la explotación sexual. En cuanto a los Nicks elegidos por los menores, la elección del Nick por parte del menor puede dar pistas sobre su estado emocional o fragilidad, por ejemplo hemos advertido nombres de posibles víctimas como SadBoy, SweetGirl, Princexita, etc. Respecto al número de amistades y número de publicaciones/fotos del menor, podemos concluir que un menor con un número pequeño es blanco para un presunto pedófilo, pues no es lo habitual en las redes sociales por parte de niños y adolescentes, donde para ellos lo importante es disponer de muchas amistades, que de alguna forma les podría dar cierto prestigio de fama. Por lo tanto, un menor con un número pequeño de amistades, es susceptible de carencias afectivas o/y problemas de alguna índole. Por último, del estudio del perfil del pedófilo en las redes sociales, y para que el presunto pedófilo consiga un fácil acercamiento al menor, éste hace uso calcado de sus mismos gustos: músicas, aficiones, deportes, cine, etc. Esto lo hemos operativizado en la Variable Registro de Actividad (RA)

Una vez realizado el análisis del perfil de **Víctima** a nivel conceptual, se procede al desarrollo e implementación de dicho perfil en nuestro Modelo de Riesgos, de forma que las variables operativizadas serían:

- **Variable SITUACIÓN FAMILIAR:**

- **Conflicto Familiar:** se refiere a la vulnerabilidad a la que el menor pudiera estar sometido/a como consecuencia de la situación familiar en la que se encuentra inmerso. El presunto pedófilo elige sus víctimas atendiendo a esta característica, de forma que la falta de estabilidad familiar del menor puede influir en la decisión de elección del agresor (junto con otras variables), así, por ejemplo, esta falta de estabilidad podría referirse al hecho de que el menor viva inmerso en una situación de separación por parte de los padres, o que viva en un entorno de conflictos familiares como abusos, peleas, etc. En nuestro Modelo de Riesgos, se interpretaría como Ausencia/Presencia de Conflicto Familiar.
- **Domicilio:** Para el presunto agresor es preferible elegir víctimas que se encuentren geográficamente más o menos cerca pues un objetivo final podría ser el contacto en persona con el menor.

- **Variable INFORMACIÓN SENSIBLE DEL MENOR:**

- **Edad:** Definida como riesgo para nuestro Modelo para menores o iguales de 13 años.
- **Género:** Definida como riesgo para género masculino
- **Nick:** Se define asociado a riesgo cuando su significado pueda tener una connotación que alerte del estado emocional del menor
- **Nº amistades (A):** Consideramos que si A es menor 30 está asociado a riesgo.

- **Nº publicaciones/Fotos (NP):** Si P es menor o igual de 10 lo asociamos a riesgo, porque aunque se podría pensar que cuantas más publicaciones o fotos más fácil es para el presunto pedófilo elegir a su víctima, por sí sola no es predictiva, porque al igual que ocurre con número de amistades, es el menor con un número escaso o no muy grande de publicaciones el no habitual en las redes sociales, y el que puede ser vulnerable a su ataque por carencias de tipo afectivas.
- **Registro Actividad (RA):** Se asocia riesgo si RA del menor presenta coincidencias con el del presunto pedófilo

Todos estos parámetros se han dicotomizado con un **1** para riesgo y un **0** para no riesgo, y los valores asignados como límites pueden ser modificados atendiendo a otras necesidades. El Modelo de Riesgos del Perfil Víctima tendría la siguiente forma:

Figura 8. Variables que definen el perfil de pedófilo en el Modelo de Riesgos

A.1 PERFIL VICTIMA					TOTAL	5
A.1.1 SITUACIÓN FAMILIAR					TOTAL	1
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO	
A.1.1.1	CONFLICTO FAMILIAR	0	AUSENCIA	ESTABILIDAD FAMILIAR		
A.1.1.2	DOMICILIO	1	CERCA	MISMA CIUDAD/ PROVINCIA	<div></div>	
TOTAL		1				
A.1.2 INFORMACIÓN SENSIBLE VÍCTIMA					TOTAL	4
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO	
A.1.2.1	EDAD	1	E <= 13	IGUAL O MENOR DE 13 AÑOS	<div></div>	
A.1.2.2	GÉNERO	1	M	MASCULINO	<div></div>	
A.1.2.3	NICK	1	SIGNIFICADO RIESGO	FRAGILIDAD EMOCIONAL	<div></div>	
A.1.2.4	Nº AMISTADES (A)	0	A >= 30	Nº DE AMISTADES IGUAL O SUPERIOR A 30		
A.1.2.5	Nº PUBLICACIONES/FOTOS (PF)	0	PF>10	Nº DE PUBLICACIONES O FOTOS SUPERIOR A 10		
A.1.2.6	REGISTRO ACTIVIDAD (RA)	1	RA = RA PEDÓFILO	COINCIDENCIA EN AFICIONES, MÚSICA, DEPORTE, ETC.	<div></div>	
TOTAL		4				

Fuente: Elaboración propia

3.2.2 Perfil agresor sexual

¿Qué ocurre en la mente de un pedófilo?

Expertos, como E. Echeburúa, catedrático de Psicología Clínica en la Universidad del País Vasco [20], como hemos comentado anteriormente en las posibles causas, no han dado aún con una respuesta clara para explicar esta patología, la pedofilia. Los pedófilos sienten un impulso irrefrenable y saltan de la fantasía a la realidad. Y precisamente Internet, con sus comunidades virtuales y foros, les han dado alas y han reforzado su engaño.

¿Cómo se distingue a un pedófilo? ¿Cómo operan los pedófilos?

Generalmente, acostumbran a crear distintos perfiles, obviamente con nombres falsos, simulando tener la misma edad de sus potenciales víctimas. Además, recopilan la mayor información de sus posibles víctimas a partir de los datos que éstas comparten a través de las redes sociales. Para conseguir el objetivo, el pedófilo sigue un proceso denominado como GROOMING, y que refiere con un tipo de patrón de comportamiento dividido en tres etapas con una clara secuencialidad, y cuyas etapas son:

- **Fase de contacto (Gaining Access)**

Fase de inicio donde el presunto agresor se centra en la captación del menor. Para ello, hará uso de toda la información del menor de las redes sociales para parecer tener muchas cosas en común, y así ganarse poco a poco la confianza del menor.

- **Fase de toma de confianza (Deceptive Relationship)**

Una vez habiendo sido capaz el presunto pedófilo de ganarse su confianza, ahora debe establecer una relación más estrecha, donde hará uso de la información emocional del menor para que éste se sienta comprendido, e incluso atendido en sus carencias afectivas.

- **Fase Sexual (Sexual Affair)**

Fase final donde el pedófilo alcanza su objetivo final, intención sexual, mediada por envío de fotografías, vídeos, y demás.

La siguiente figura (Figura 9), elaborada por Rick Braschler, Six Stages of Child Grooming, representa gráficamente el proceso llevado a cabo por los pedófilos [21].

Figura 9. Los seis estado del Child-Grooming



Fuente: Braschler, R. (2014). Six Stages of Child Grooming – Abuse prevention, detection, response. Recuperado el 14 de marzo de 2016 de <http://www.camphow.com/camping/risk-management/stages-child-grooming-abuse-prevention-detection-response/trackback>

Una vez siendo conocedores del proceso llevado a cabo por los agresores sexuales, no nos extraña la empatía y simpatía para con el menor. Para este complejo proceso, metódicamente estudiado y definido, los pedófilos también utilizan unas marcas distintivas en las redes, que van desde la elección de su Nick, con connotaciones de dominio e incluso sexual, hasta el uso de determinados símbolos, como es el caso de los triángulos y corazones. Los símbolos son siempre compuestos por la unión de 2 similares, uno dentro del otro. El de forma mayor identifica al adulto, la menor al niño. La diferencia de tamaños entre ellos muestra una preferencia por niños mayores o menores en cuanto a la edad. Los hombres son triángulos, corazones el de las mujeres. Los símbolos se encuentran en sitios como, monedas, joyas, anillos, colgantes, etc., entre otros objetos. Los triángulos representan a los hombres que le gustan los niños (el detalle cruel es el triángulo más pequeño, que representa al hombre que le gustan los niños bien pequeños); el corazón son hombres (o mujeres) que gustan de niñas y la mariposa representa a quienes gustan de ambos, La idea de los triángulos concéntricos es la de una figura mayor envolviendo a una figura menor, dentro de una genialidad perversa de un concepto gráfico. Existe un nivel mayor de crueldad, porque a esos seres les gusta exhibirse en códigos para otros, haciendo eso símbolos en bisutería, moneda, trofeos, adhesivos, etc. Lamentablemente, es el diseño gráfico al servicio del trastorno.

Éstos son algunos de los símbolos que usan los Pedófilos:

Figura 10. Símbolos usados por pedófilos



Por todo lo expuesto anteriormente, se decide implementar en nuestro Modelo de Riesgos una variable que controle el uso de este tipo de símbolos y también se apuesta por un control preventivo del riesgo teniendo en cuenta el significado y verificación del Nick, donde además de hacer un “check” a una posible connotación de riesgo del significado, también se verifica si pertenece a alguno utilizado por los pedófilos de los que se tiene en bases de datos, en nuestro caso, hemos usado como Base de Datos de referencia www.perverted-justice.com. Además se verifica si la fotografía del perfil del presunto pedófilo es utilizada en cualquier otro perfil. Se dispone de varios portales especializados en un sistema de detección de fraude de identidad con alta precisión.

Por otra parte, tras un minucioso estudio de conversaciones de pedófilos de www.perverted-justice.com, hemos concluido con otros factores de riesgo, como son el promedio de conexión con la víctima y el número de peticiones de conexión. En primer lugar, resulta lógico, tras la explicación del proceso de GROOMING, que el agresor sexual necesita su tiempo para “captar” a su víctima. Y además, se concluye que aunque el presunto pedófilo es el que realiza la primera conexión, lógicamente pues es el que elige a su víctima, en las posteriores conexiones, es el menor el que se encarga de la mayoría de inicios de conversaciones, cosa razonable después de una toma de contacto y toma de confianza; el pedófilo no se puede arriesgar a que el menor pueda estar acompañado de un adulto.

Una vez realizado el análisis del perfil del Pedófilo a nivel conceptual, se procede al desarrollo e implementación de dicho perfil en nuestro Modelo de Riesgos, de forma que las variables operativizadas serían:

- **IDENTIFICACIÓN**
 - **Significado Nick:** Se asignará riesgo cuando el significado tenga una connotación sexual o dominante.
 - **Verificación Nick:** Se asignará riesgo cuando sea localizado en algún fichero de datos de conversaciones pedófilas.
 - **Identificación foto perfil:** Se asignará riesgo cuando sea verificada como identidad compartida con otro usuario.
 - **Símbolo:** Se asignará riesgo cuando pertenezca al tipo de símbolos usados por los pedófilos en las redes sociales.
- **TIEMPO PROMEDIO CONEXIÓN CON VÍCTIMA:** Variable con valor 1 de riesgo si el tiempo promedio de conexión es mayor o igual a media hora.
- **NÚMERO DE PETICIONES CONEXIÓN A LA VÍCTIMA:** Variable con valor 1 de riesgo cuando es menor de 2.

Todos estos parámetros han sido dicotomizados, con un 1 para riesgo y un 0 para no riesgo, y los valores asignados como límites pueden ser modificados atendiendo a otras necesidades. El Modelo de Riesgos del Perfil Pedófilo tendría la siguiente forma:

Figura 11. Variables que definen el perfil de pedófilo en el Modelo de Riesgos

A.2 PERFIL PEDÓFILO						TOTAL	6
A.2.1 IDENTIFICACIÓN						TOTAL	4
CÓD. VARIABL	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO		
A.2.1.1	SIGNIFICADO NICK	1	NICK SIGNIFICATIVO	SIGNIFICADO SEXUAL O DOMINANTE DEL NICK			
A.2.1.2	VERIFICACIÓN NICK	1	NICK VERIFICADO	SE ENCUENTRA EN REGISTRO			
A.2.1.3	IDENTIFICACIÓN FOTO	1	CHECK FOTO	LA FOTO PERTENECE A OTRO PERFIL DE OTRO USUARIO			
A.2.1.4	SÍMBOLO	1	SÍMBOLO ASOCIADO A PEDOFILIA	SÍMBOLO ASOCIADO A PEDOFILIA			
TOTAL		4					
A.2.2 TIEMPO CONEXIÓN						TOTAL	1
CÓD. VARIABL	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO		
A.2.2.1	TIEMPO PROMEDIO CONEXIÓN CON VÍCTIMA	1	TP > 1/2 HORA	SUPERIOR A MEDIA HORA			
TOTAL		1					
A.2.3 Nº PETICIONES CONEXIÓN						TOTAL	1
CÓD. VARIABL	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO		
A.2.3.1	Nº PETICIONES CONEXIÓN A LA VÍCTIMA	1	PC <=2	INFERIOR O IGUAL A DOS SOLICITUDES			
TOTAL		1					

Fuente: Elaboración propia

3.2.3 Perfil Conversación

El acto de diálogo o conversación es la variable que relaciona al presunto pedófilo con la víctima, y que se considera como variable definitoria junto con la de víctima y agresor para la detección de contenido pedófilo. Es por ello, que esta variable ha sido estructurada atendiendo a sus características físicas en sí mismas como pueden ser la densidad conversacional, duración y horario de la conversación; atendiendo a un análisis de contenido: léxico, semántico y sintáctico; y finalmente atendiendo a la etapa de Grooming en la que podría encontrarse la víctima y agresor.

Las características físicas de la conversación han sido definidas en la densidad conversacional, duración y horario. Se define *la unidad conversacional* como la unidad mínima de conversación establecida entre un emisor y un receptor, es decir, en nuestro caso entre la víctima y el pedófilo. Una vez definida, se conceptualizara la densidad conversacional como el número de unidades conversacionales intercambiadas entre víctima y agresor en cada conexión. Revisadas conversaciones de pedófilos suele ser un número considerable. La duración del proceso también es una variable relevante en nuestra predicción pues el proceso Grooming no es demasiado largo, el pedófilo quiere conseguir el objetivo lo antes posible, pero paso a paso. Y en cuanto al horario, es lógico pensar que el pedófilo use un horario donde considere que el menor está solo, o poco vigilado: en horario de mañana están en colegio, y a partir de la salida del colegio hasta la posible llegada de sus padres del trabajo, sería la franja horaria de actuación del pedófilo.

En relación a la variable de análisis de contenido, hemos decidido hacer un estudio a nivel léxico, semántico y sintáctico. El primer nivel de análisis atiende al nivel léxico de las palabras utilizadas en la conversación. Es importante asegurarse de que la conversación tiene lugar

niños, y sabemos que éstos utilizan una forma muy específica de lenguaje que incluye desde gran cantidad de *emoticonos*, *abreviaturas* y *un lenguaje particular pre-adolescente*. Todos estos términos pueden ser categorizados mediante sistemas de procesamiento y análisis léxico, mencionados en el apartado de Estado del Arte. En cuanto al nivel semántico, podríamos hacer un desglose atendiendo a si la carga del contenido es emocional, de acoso o sexual. Quizás sea la carga con contenido sexual la mejor variable predictiva de detección de pedófilos, a nivel semántico, y la que de una manera contundente debe ser controlada y definida como alto riesgo. Los contenidos con cargas emocionales, e incluso de acoso son más difíciles de detectar. También existen sistemas de procesamiento para categorizar palabras atendiendo al nivel semántico (LIWC). Y por último, el nivel sintáctico atendería a la estructura de la frase, por lo que esta variable se operativizará con su presencia o ausencia en el Corpus disponible en www.perverted-justice.com.

En cuanto a último nivel de estudio de la variable Conversación, se decide introducir el concepto de etapas de Grooming: tres etapas, que por otra parte tienen un contenido distinto a nivel características físicas conversacionales y nivel de análisis de contenido. La complejidad de dicha asociación, nos hace definir como etapa de máximo riesgo, con una probabilidad alta, la etapa sexual (tercera etapa), absolutamente inaceptable, que vendría asociada a densidades de conversaciones altas, horarios marcados, y una duración no muy elevada en días y alto contenido a nivel semántico con clara carga sexual.

Una vez realizado el análisis del perfil de la Conversación a nivel conceptual, se procede al desarrollo e implementación de dicho perfil en nuestro Modelo de Riesgos, de forma que las variables operativizadas serían:

- **CARACTERÍSTICAS FÍSICAS DE LA CONVERSACIÓN**

- **Densidad conversacional:** Se asume como Riesgo una densidad mayor o igual a 10 unidades conversacionales.
- **Duración:** Se asume como Riesgo una duración menor o igual a 7 días
- **Horario:** Se asume como Riesgo un horario a partir de las 13:00 horas.

- **ANÁLISIS DE CONTENIDO**

- **Léxico –Emoticonos:** Si el número de emoticonos por unidad conversacional es superior a 5 supone un riesgo
- **Léxico-Abreviaturas:** Si el número de abreviaturas por unidad conversacional es superior a 10 supone riesgo
- **Léxico-Lenguaje_preadolescente:** La presencia de dicho lenguaje supone riesgo
- **Semántico:** La presencia de términos con carga sexual supone riesgo
- **Sintáctico:** La presencia de las frases en el Corpus de la BD de pedófilos

- ETAPAS

- Sexual: Riesgo Máximo

El modelo de Riesgos para el perfil de la Variable Conversación sería:

Figura 12. Variables que definen el perfil de conversación en el Modelo de Riesgos

A.3 CONVERSACION						TOTAL	2
A.3.1 CARACTERÍSTICAS FÍSICAS CONVERSACIÓN						TOTAL	2
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO		
A.3.1.1	DENSIDAD CONVERSACIONAL	1	Nº UNIDADES CONVERSACIONALES >= 1	DENSIDAD CONVERSACIONAL ALTA			
A.3.1.2	DURACIÓN	1	Nº DÍAS <7	CORTA DURACIÓN			
A.3.1.3	HORARIO	0	HASTA LAS 13:00 HORAS	HORARIO DE MAÑANA (A.M.)			
TOTAL		2					
A.3.2 ANÁLISIS DE CONTENIDOS						TOTAL	0
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO		
A.3.2.1	LÉXICO: EMOTICONOS	0	EMOTICONOS <5	BAJO NÚMERO DE EMOTICONOS			
A.3.2.2	LÉXICO: ABREVIATURAS	0	ABREVIATURAS <10	BAJO NÚMERO DE ABREVIATURAS			
A.3.2.3	LÉXICO: LENGUAJE PREADOLESCENTE	0	AUSENCIA	AUSENCIA LENGUAJE PREADOLESCENTE			
A.3.2.4	SEMÁNTICO: TÉRMINOS CON CARGA SEXUAL	0	AUSENCIA	AUSENCIA TÉRMINOS CON CARGA SEXUAL			
A.3.2.5	SINTÁCTICO: CORPUS EN BASE DE DATOS	0	AUSENCIA	AUSENCIA CORPUS EN BASE DE DATOS			
TOTAL		0					
A.3.3 ETAPAS CONSOLIDADAS						TOTAL	0
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO		
A.3.3.1	ETAPA EN LA QUE ENCUENTRA	0	ETAPA 1 O 2	ETAPAS CONTACTO - GROOMING			
TOTAL		0					

Fuente: Elaboración propia

3.3.Las TICs al Servicio de la Detección de Conversaciones de Riesgo Pedófilo: Propuesta de un Modelo de Riesgos

A continuación mostramos el resultado de la incorporación e implementación en el Modelo de Riesgos (Figura 13) de la matriz de variables tras su Operativización (Figura 14).

Figura 13. Modelo de Riesgos

					TOTAL RIESGO	12	52%
A.1 PERFIL VÍCTIMA					TOTAL	5	
A.1.1 SITUACIÓN FAMILIAR					TOTAL	1	
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO		
A.1.1.1	CONFLICTO FAMILIAR	0	AUSENCIA	ESTABILIDAD FAMILIAR			
A.1.1.2	DOMICILIO	1	CERCA	MISMA CIUDAD/PROVINCIA			
TOTAL		1					
A.1.2 INFORMACIÓN SENSIBLE VÍCTIMA					TOTAL	4	
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO		
A.1.2.1	EDAD	1	E <= 13	IGUAL O MENOR DE 13 AÑOS			
A.1.2.2	GÉNERO	1	M	MASCULINO			
A.1.2.3	NICK	1	SIGNIFICADO RIESGO	FRAGILIDAD EMOCIONAL			
A.1.2.4	Nº AMISTADES (A)	0	A >= 30	Nº DE AMISTADES IGUAL O SUPERIOR A 30			
A.1.2.5	Nº PUBLICACIONES/FOTOS (PF)	0	PF>10	Nº DE PUBLICACIONES O FOTOS SUPERIOR A 10			
A.1.2.6	REGISTRO ACTIVIDAD (RA)	1	RA = RA PEDÓFILO	COINCIDENCIA EN AFICIONES, MÚSICA, DEPORTE, ETC.			
TOTAL		4					
A.2 PERFIL PEDÓFILO					TOTAL	5	
A.2.1 IDENTIFICACIÓN					TOTAL	3	
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO		
A.2.1.1	SIGNIFICADO NICK	1	NICK SIGNIFICATIVO	SIGNIFICADO SEXUAL O DOMINANTE DEL NICK			
A.2.1.2	VERIFICACIÓN NICK	1	NICK VERIFICADO	SE ENCUENTRA EN REGISTRO			
A.2.1.3	IDENTIFICACIÓN FOTO	0	NO CHECK FOTO	LA FOTO NO PERTENECE A OTRO PERFIL DE OTRO USUAR			
A.2.1.4	SÍMBOLO	1	SÍMBOLO ASOCIADO A PEDOFILIA	SÍMBOLO ASOCIADO A PEDOFILIA			
TOTAL		3					
A.2.2 TIEMPO CONEXIÓN					TOTAL	1	
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO		
A.2.2.1	TIEMPO PROMEDIO CONEXIÓN CON VÍCTIMA	1	TP > 1/2 HORA	SUPERIOR A MEDIA HORA			
TOTAL		1					
A.2.3 Nº PETICIONES CONEXIÓN					TOTAL	1	
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO		
A.2.3.1	Nº PETICIONES CONEXIÓN A LA VÍCTIMA	1	PC <=2	INFERIOR O IGUAL A DOS SOLICITUDES			
TOTAL		1					
A.3 CONVERSACION					TOTAL	2	
A.3.1 CARACTERÍSTICAS FÍSICAS CONVERSACIÓN					TOTAL	2	
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO		
A.3.1.1	DENSIDAD CONVERSACIONAL	1	Nº UNIDADES CONVERSACIONALES >= 1	DENSIDAD CONVERSACIONAL ALTA			
A.3.1.2	DURACIÓN	1	Nº DÍAS <7	CORTA DURACIÓN			
A.3.1.3	HORARIO	0	HASTA LAS 13:00 HORAS	HORARIO DE MAÑANA (A.M.)			
TOTAL		2					
A.3.2 ANÁLISIS DE CONTENIDOS					TOTAL	0	
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO		
A.3.2.1	LÉXICO: EMOTICONOS	0	EMOTICONOS <5	BAJO NÚMERO DE EMOTICONOS			
A.3.2.2	LÉXICO: ABREVIATURAS	0	ABREVIATURAS <10	BAJO NÚMERO DE ABREVIATURAS			
A.3.2.3	LÉXICO: LENGUAJE PREADOLESCENTE	0	AUSENCIA	AUSENCIA LENGUAJE PREADOLESCENTE			
A.3.2.4	SEMÁNTICO: TÉRMINOS CON CARGA SEXUAL	0	AUSENCIA	AUSENCIA TÉRMINOS CON CARGA SEXUAL			
A.3.2.5	SINTÁCTICO: CORPUS EN BASE DE DATOS	0	AUSENCIA	AUSENCIA CORPUS EN BASE DE DATOS			
TOTAL		0					
A.3.3 ETAPAS CONSOLIDADAS					TOTAL	0	
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO		
A.3.3.1	ETAPA EN LA QUE ENCUENTRA	0	ETAPA 1 O 2	ETAPAS CONTACTO - GROOMING			
TOTAL		0					

Fuente: Elaboración propia

Figura 14. Conjunto de variables del Modelo de Riesgos

ÁMBITO GENERAL	
Cód. variable	Descripción
A.1	PERFIL VÍCTIMA
A.2	PERFIL PEDÓFILO
A.3	CONVERSACIÓN

VARIABLES VÍCTIMA					
A.1	PERFIL VÍCTIMA	A.1.1	SITUACIÓN FAMILIAR VÍCTIMA	A.1.2	INFORMACIÓN SENSIBLE VÍCTIMA
A.1.1	SITUACIÓN FAMILIAR	A.1.1.1	CONFLICTO FAMILIAR	A.1.2.1	EDAD
A.1.2	INFORMACIÓN SENSIBLE VÍCTIMA	A.1.1.2	DOMICILIO	A.1.2.2	GÉNERO
				A.1.2.3	NICK
				A.1.2.4	Nº AMISTADES (A)
				A.1.2.5	Nº PUBLICACIONES/FOTOS (PF)
				A.1.2.6	REGISTRO ACTIVIDAD (RA)

VARIABLES PEDÓFILO					
A.2	PERFIL PEDÓFILO	A.2.1	IDENTIFICACIÓN PEDÓFILO	A.2.2	TIEMPO PROMEDIO - TP
A.2.1	IDENTIFICACIÓN	A.2.1.1	SIGNIFICADO NICK	A.2.2.1	TIEMPO PROMEDIO CONEXIÓN CON VÍCTIMA
A.2.2	TIEMPO CONEXIÓN	A.2.1.2	VERIFICACIÓN NICK		
A.2.3	Nº PETICIONES CONEXIÓN	A.2.1.3	IDENTIFICACIÓN FOTO		
		A.2.1.4	SÍMBOLO		

VARIABLES CONVERSACIÓN					
A.3	CARACTERÍSTICAS CONVERSACIÓN	A.3.1	CARACTERÍSTICAS FÍSICAS CONVERSACIÓN	A.3.2	ANÁLISIS DE CONTENIDO
A.3.1	CARACTERÍSTICAS FÍSICAS CONVERSACIÓN	A.3.1.1	DENSIDAD CONVERSACIONAL	A.3.2.1	LÉXICO: EMOTICONOS
A.3.2	ANÁLISIS DE CONTENIDOS	A.3.1.2	DURACIÓN	A.3.2.2	LÉXICO: ABREVIATURAS
A.3.3	ETAPAS CONSOLIDADAS	A.3.1.3	HORARIO	A.3.2.3	LÉXICO: LENGUAJE PREADOLESCENTE
				A.3.2.4	SEMÁNTICO: TÉRMINOS CON CARGA SEXUAL
				A.3.2.5	SINTÁCTICO: CORPUS EN BASE DE DATOS

A.3.3	ETAPAS CONSOLIDADAS
A.3.3.1	ETAPA EN LA QUE ENCUENTRA

Fuente: Elaboración propia

3.4. Implementación en desarrollo con Técnicas SVMs

Las máquinas de soporte vectorial, máquinas de vectores de soporte o máquinas de vector soporte (Support Vector Machines, SVMs) son un conjunto de algoritmos de aprendizaje supervisado que están relacionados con problemas de clasificación y regresión donde, dado un conjunto de puntos, subconjunto de un conjunto mayor (espacio), en el que cada uno de ellos pertenece a una de dos posibles categorías, un algoritmo basado en SVM construye un modelo capaz de predecir si un punto nuevo (cuya categoría desconocemos) pertenece a una categoría o a la otra.

Actualmente el problema del reconocimiento de objetos se puede resolver con otras técnicas basadas, por ejemplo, en redes neuronales, sin embargo, las SVM son una alternativa que en muchos campos se han mostrado superiores, como en el reconocimiento de datos biométricos (voz, reconocimiento de retina, facial...). Bajo este concepto la futura implementación en desarrollo de nuestro proyecto haría uso de estas técnicas.

Como ejemplo de aplicación de técnicas SVM a nuestro contexto, existen trabajos de investigación del uso de técnicas biométricas basadas en rasgos faciales para la detección de rangos de edad, con el propósito de evitar que adultos se hagan pasar por menores, o incluso que ciertos menores puedan acceder a redes sociales, cuyo acceso debe estar restringido por su edad. Los resultados muestran que es posible hacer esta distinción entre adultos y menores, seleccionando edades umbrales cercanas a los 18 años, con tasas de acierto cercanas al 80 %, y empleando clasificadores basados en máquinas de vector soporte (SVMs).

Capítulo IV

Validación práctica

Para la comprobación del correcto funcionamiento del Modelo de Riesgos partimos de un ejemplo práctico de una unidad conversacional que sometemos a análisis contra las variables de los perfiles del modelo. Con este fin se ha elegido un fragmento de conversación procedente de www.perverted-justice.com.

4.1. Validación Modelo de Riesgos

Partiendo de una hipotética unidad conversacional entre el presunto pedófilo y su víctima, los datos asignados a las variables de nuestros perfiles podrían tener el siguiente aspecto:

- **PERFIL VÍCTIMA**

Conflicto familiar: NO **Domicilio:** misma ciudad

Edad: 12años **Género:** masculino **Nick=**Sadgirl; **A=**35 **NP=**11 **RA=** 1

- **PERFIL PRESUNTO PEDÓFILO**

Significado Nick=Capitán35 **Verificación Nick=**SI; **Identificación Foto=**0; **Símbolo=**1

Tiempo Promedio Conexión=3 horas;

Número Peticiones Conexión=2

- **PERFIL CONVERSACIÓN**

Densidad Conversacional=30 **Duración=**3 **Horario=**15:00

Emoticonos=7 **Abreviaturas=**8 **Lenguaje** **adolescente:**SI

Términos contenido sexual: SI **Corpus:** NO **Etapas sexual:** NO

La anterior información se presenta resumida en la Tabla 1:

Tabla 1. Ejemplo de codificación de variables en el Modelo de Riesgos

VARIABLES DEL MODELO	CONTENIDO	CODIFICACIÓN
Conflicto familiar	NO	0
Domicilio	SI	1
Edad	SI	1
Género	SI	1
Nick	Sadgirl	1
Amistades	35	0
Nº Publicaciones fotos	11	0
Registro actividad	SI	1
Significado Nick	Capitán 35	1
Verificación Nick	SI	1
Identificación de la foto	NO	0
Símbolo	SI	1
Tiempo promedio Conexión	3 HORAS	1
Nº peticiones de conexión	2 PETICIONES	1
Densidad conversacional	30 UNIDADES	1
Duración	3 DÍAS	1
Horario	15:00 horas	1
Emoticonos	7	1
Abreviaturas	8	0
Lenguaje adolescente	Si	1
Términos con contenido sexual	SI	1
Corpus	NO	0
Etapas sexual	NO	0

Fuente: Elaboración propia

A partir de los datos de nuestro ejemplo sacado de una hipotética unidad conversacional mantenida por el presunto pedófilo y la víctima, nuestro vector de variables predictivas de detección de conversaciones con contenido pedófilo se mostraría de la siguiente manera:

(0,1,1,1,0,0,1,1,1,0,1,1,1,1,1,0,1,1,0,0)

Por lo que partiendo de las matrices iniciales del Modelo de Riesgos donde se asume la existencia/no existencia de riesgos, que ofrecerían el siguiente aspecto (Figura 15 y Figura 16).

- MATRICES SIN RIESGO Y CON RIESGO MÁXIMO**

Figura 15. Matriz de variables sin riesgo

					TOTAL RIESGO	0	0%
A.1 PERFIL VÍCTIMA					TOTAL	0	
A.1.1 SITUACIÓN FAMILIAR					TOTAL	0	
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO		
A.1.1.1	CONFLICTO FAMILIAR	0	AUSENCIA	ESTABILIDAD FAMILIAR			
A.1.1.2	DOMICILIO	0	LEJOS	DISTINTA CIUDAD/PROVINCIA			
		TOTAL	0				
A.1.2 INFORMACIÓN SENSIBLE VÍCTIMA					TOTAL	0	
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO		
A.1.2.1	EDAD	0	E > 13	SUPERIOR A 13 AÑOS			
A.1.2.2	GÉNERO	0	F	FEMENINO			
A.1.2.3	NICK	0	NO RIESGO	NO FRAGILIDAD EMOCIONAL			
A.1.2.4	Nº AMISTADES (A)	0	A >= 30	Nº DE AMISTADES IGUAL O SUPERIOR A 30			
A.1.2.5	Nº PUBLICACIONES/FOTOS (PF)	0	PF>10	Nº DE PUBLICACIONES O FOTOS SUPERIOR A 10			
A.1.2.6	REGISTRO ACTIVIDAD (RA)	0	NO COINCIDENCIA	NO COINCIDENCIA EN AFICIONES, MÚSICA, DEPORTE, ETC			
		TOTAL	0				
A.2 PERFIL PEDOFILO					TOTAL	0	
A.2.1 IDENTIFICACIÓN					TOTAL	0	
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO		
A.2.1.1	SIGNIFICADO NICK	0	NICK NO SIGNIFICATIVO	NO SIGNIFICADO SEXUAL O DOMINANTE DEL NICK			
A.2.1.2	VERIFICACIÓN NICK	0	NICK NO VERIFICABLE	NO SE ENCUENTRA EN REGISTRO			
A.2.1.3	IDENTIFICACIÓN FOTO	0	NO CHECK FOTO	LA FOTO NO PERTENECE A OTRO PERFIL DE OTRO USUARIO			
A.2.1.4	SÍMBOLO	0	NO SÍMBOLO	SÍMBOLO NO ASOCIADO A PEDOFILIA			
		TOTAL	0				
A.2.2 TIEMPO CONEXIÓN					TOTAL	0	
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO		
A.2.2.1	TIEMPO PROMEDIO CONEXIÓN CON VÍCTIMA	0	TP <= MEDIA HORA	IGUAL O INFERIOR A MEDIA HORA			
		TOTAL	0				
A.2.3 Nº PETICIONES CONEXIÓN					TOTAL	0	
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO		
A.2.3.1	Nº PETICIONES CONEXIÓN A LA VÍCTIMA	0	PC >2	SUPERIOR A DOS SOLICITUDES			
		TOTAL	0				
A.3 CONVERSACION					TOTAL	0	
A.3.1 CARACTERÍSTICAS FÍSICAS CONVERSACIÓN					TOTAL	0	
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO		
A.3.1.1	DENSIDAD CONVERSACIONAL	0	Nº UNIDADES CONVERSACIONALES <10	DENSIDAD CONVERSACIONAL BAJA			
A.3.1.2	DURACIÓN	0	Nº DÍAS >= 7	LARGA DURACIÓN			
A.3.1.3	HORARIO	0	HASTA LAS 13:00 HORAS	HORARIO DE MAÑANA (A.M.)			
		TOTAL	0				
A.3.2 ANÁLISIS DE CONTENIDOS					TOTAL	0	
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO		
A.3.2.1	LÉXICO: EMOTICONOS	0	EMOTICONOS <5	BAJO NÚMERO DE EMOTICONOS			
A.3.2.2	LÉXICO: ABREVIATURAS	0	ABREVIATURAS <10	BAJO NÚMERO DE ABREVIATURAS			
A.3.2.3	LÉXICO: LENGUAJE PREADOLESCENTE	0	AUSENCIA	AUSENCIA LENGUAJE PREADOLESCENTE			
A.3.2.4	SEMÁNTICO: TÉRMINOS CON CARGA SEXUAL	0	AUSENCIA	AUSENCIA TÉRMINOS CON CARGA SEXUAL			
A.3.2.5	SINTÁCTICO: CORPUS EN BASE DE DATOS	0	AUSENCIA	AUSENCIA CORPUS EN BASE DE DATOS			
		TOTAL	0				
A.3.3 ETAPAS CONSOLIDADAS					TOTAL	0	
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO		
A.3.3.1	ETAPA EN LA QUE ENCUENTRA	0	ETAPA 1 O 2	ETAPAS CONTACTO - GROOMING			
		TOTAL	0				

Fuente: Elaboración propia

Figura 16. Matriz de variables con riesgo máximo

TOTAL RIESGO					23	100%
A.1 PERFIL VÍCTIMA					TOTAL	8
A.1.1 SITUACIÓN FAMILIAR					TOTAL	2
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO	
A.1.1.1	CONFLICTO FAMILIAR	1	PRESENCIA	FALTA ESTABILIDAD FAMILIAR	■	
A.1.1.2	DOMICILIO	1	CERCA	MISMA CIUDAD/PROVINCIA	■	
TOTAL		2				
A.1.2 INFORMACIÓN SENSIBLE VÍCTIMA					TOTAL	6
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO	
A.1.2.1	EDAD	1	E <= 13	IGUAL O MENOR DE 13 AÑOS	■	
A.1.2.2	GÉNERO	1	M	MASCULINO	■	
A.1.2.3	NICK	1	SIGNIFICADO RIESGO	FRAGILIDAD EMOCIONAL	■■	
A.1.2.4	Nº AMISTADES (A)	1	A < 30	Nº DE AMISTADES INFERIOR A 30	■■■	
A.1.2.5	Nº PUBLICACIONES/FOTOS (PF)	1	PF<=10	Nº DE PUBLICACIONES O FOTOS INFERIOR A 10	■■■■	
A.1.2.6	REGISTRO ACTIVIDAD (RA)	1	RA = RA PEDÓFILO	COINCIDENCIA EN AFICCIONES, MÚSICA, DEPORTE, ETC.	■■■■■	
TOTAL		6				
A.2 PERFIL PEDÓFILO					TOTAL	6
A.2.1 IDENTIFICACIÓN					TOTAL	4
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO	
A.2.1.1	SIGNIFICADO NICK	1	NICK SIGNIFICATIVO	SIGNIFICADO SEXUAL O DOMINANTE DEL NICK	■	
A.2.1.2	VERIFICACIÓN NICK	1	NICK VERIFICADO	SE ENCUENTRA EN REGISTRO	■	
A.2.1.3	IDENTIFICACIÓN FOTO	1	CHECK FOTO	LA FOTO PERTENECE A OTRO PERFIL DE OTRO USUARIO	■■	
A.2.1.4	SÍMBOLO	1	SÍMBOLO ASOCIADO A PEDOFILIA	SÍMBOLO ASOCIADO A PEDOFILIA	■■■	
TOTAL		4				
A.2.2 TIEMPO CONEXIÓN					TOTAL	1
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO	
A.2.2.1	TIEMPO PROMEDIO CONEXIÓN CON VÍCTIMA	1	TP > 1/2 HORA	SUPERIOR A MEDIA HORA	■	
TOTAL		1				
A.2.3 Nº PETICIONES CONEXIÓN					TOTAL	1
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO	
A.2.3.1	Nº PETICIONES CONEXIÓN A LA VÍCTIMA	1	PC <=2	INFERIOR O IGUAL A DOS SOLICITUDES	■	
TOTAL		1				
A.3 CONVERSACION					TOTAL	9
A.3.1 CARACTERÍSTICAS FÍSICAS CONVERSACIÓN					TOTAL	3
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO	
A.3.1.1	DENSIDAD CONVERSACIONAL	1	Nº UNIDADES CONVERSACIONALES >= 1	DENSIDAD CONVERSACIONAL ALTA	■	
A.3.1.2	DURACIÓN	1	Nº DÍAS <7	CORTA DURACIÓN	■	
A.3.1.3	HORARIO	1	A PARTIR DE LAS 13:00 HORAS	HORARIO DE TARDE (P.M.)	■■	
TOTAL		3				
A.3.2 ANÁLISIS DE CONTENIDOS					TOTAL	5
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO	
A.3.2.1	LÉXICO: EMOTICONOS	1	EMOTICONOS >=5	ALTO NÚMERO DE EMOTICONOS	■	
A.3.2.2	LÉXICO: ABREVIATURAS	1	ABREVIATURAS >=10	ALTO NÚMERO DE ABREVIATURAS	■	
A.3.2.3	LÉXICO: LENGUAJE PREADOLESCENTE	1	PRESENCIA	PRESENCIA LENGUAJE PREADOLESCENTE	■■	
A.3.2.4	SEMÁNTICO: TÉRMINOS CON CARGA SEXUAL	1	PRESENCIA	PRESENCIA TÉRMINOS CON CARGA SEXUAL	■■■	
A.3.2.5	SINTÁCTICO: CORPUS EN BASE DE DATOS	1	PRESENCIA	PRESENCIA CORPUS EN BASE DE DATOS	■■■■	
TOTAL		5				
A.3.3					1	TOTAL 1
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO	
A.3.3.1	ETAPA EN LA QUE ENCUENTRA	1	ETAPA 3	ETAPA SEXUAL	■	
TOTAL		1				

Fuente: Elaboración propia

Con este ejemplo, se presenta de forma sencilla un Modelo de Riesgos en la detección de conversaciones pedófilas en redes sociales, donde para ello de forma visual se ofrecen porcentajes e incluso alarmas visuales por perfiles y total, de modo que la matriz generada por nuestra entrada del vector anterior quedaría de la siguiente forma:

Figura 17. Matriz de variables según un ejemplo práctico

					TOTAL RIESGO	16	70%
A.1 PERFIL VÍCTIMA					TOTAL	5	
A.1.1 SITUACIÓN FAMILIAR					TOTAL	1	
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO		
A.1.1.1	CONFLICTO FAMILIAR	0	AUSENCIA	ESTABILIDAD FAMILIAR			
A.1.1.2	DOMICILIO	1	CERCA	MISMA CIUDAD/PROVINCIA			
		TOTAL	1				
A.1.2 INFORMACIÓN SENSIBLE VÍCTIMA					TOTAL	4	
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO		
A.1.2.1	EDAD	1	E <= 13	IGUAL O MENOR DE 13 AÑOS			
A.1.2.2	GÉNERO	1	M	MASCULINO			
A.1.2.3	NICK	1	SIGNIFICADO RIESGO	FRAGILIDAD EMOCIONAL			
A.1.2.4	Nº AMISTADES (A)	0	A >= 30	Nº DE AMISTADES IGUAL O SUPERIOR A 30			
A.1.2.5	Nº PUBLICACIONES/FOTOS (PF)	0	PF>10	Nº DE PUBLICACIONES O FOTOS SUPERIOR A 10			
A.1.2.6	REGISTRO ACTIVIDAD (RA)	1	RA = RA PEDÓFILO	COINCIDENCIA EN AFICCIONES, MÚSICA, DEPORTE, ETC.			
		TOTAL	4				
A.2 PERFIL PEDÓFILO					TOTAL	5	
A.2.1 IDENTIFICACIÓN					TOTAL	3	
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO		
A.2.1.1	SIGNIFICADO NICK	1	NICK SIGNIFICATIVO	SIGNIFICADO SEXUAL O DOMINANTE DEL NICK			
A.2.1.2	VERIFICACIÓN NICK	1	NICK VERIFICADO	SE ENCUENTRA EN REGISTRO			
A.2.1.3	IDENTIFICACIÓN FOTO	0	NO CHECK FOTO	LA FOTO NO PERTENECE A OTRO PERFIL DE OTRO USUAR			
A.2.1.4	SÍMBOLO	1	SÍMBOLO ASOCIADO A PEDOFILIA	SÍMBOLO ASOCIADO A PEDOFILIA			
		TOTAL	3				
A.2.2 TIEMPO CONEXIÓN					TOTAL	1	
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO		
A.2.2.1	TIEMPO PROMEDIO CONEXIÓN CON VÍCTIMA	1	TP > 1/2 HORA	SUPERIOR A MEDIA HORA			
		TOTAL	1				
A.2.3 Nº PETICIONES CONEXIÓN					TOTAL	1	
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO		
A.2.3.1	Nº PETICIONES CONEXIÓN A LA VÍCTIMA	1	PC <=2	INFERIOR O IGUAL A DOS SOLICITUDES			
		TOTAL	1				
A.3 CONVERSACION					TOTAL	6	
A.3.1 CARACTERÍSTICAS FÍSICAS CONVERSACIÓN					TOTAL	3	
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO		
A.3.1.1	DENSIDAD CONVERSACIONAL	1	Nº UNIDADES CONVERSACIONALES >= 1	DENSIDAD CONVERSACIONAL ALTA			
A.3.1.2	DURACIÓN	1	Nº DÍAS <7	CORTA DURACIÓN			
A.3.1.3	HORARIO	1	A PARTIR DE LAS 13:00 HORAS	HORARIO DE TARDE (P.M.)			
		TOTAL	3				
A.3.2 ANÁLISIS DE CONTENIDOS					TOTAL	3	
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO		
A.3.2.1	LÉXICO: EMOTICONOS	1	EMOTICONOS >=5	ALTO NÚMERO DE EMOTICONOS			
A.3.2.2	LÉXICO: ABREVIATURAS	0	ABREVIATURAS <10	BAJO NÚMERO DE ABREVIATURAS			
A.3.2.3	LÉXICO: LENGUAJE PREADOLESCENTE	1	PRESENCIA	PRESENCIA LENGUAJE PREADOLESCENTE			
A.3.2.4	SEMÁNTICO: TÉRMINOS CON CARGA SEXUAL	1	PRESENCIA	PRESENCIA TÉRMINOS CON CARGA SEXUAL			
A.3.2.5	SINTÁCTICO: CORPUS EN BASE DE DATOS	0	AUSENCIA	AUSENCIA CORPUS EN BASE DE DATOS			
		TOTAL	3				
A.3.3					1	TOTAL	0
CÓD. VARIABLE	DEFINICIÓN VARIABLE	VALORACIÓN	DESCRIPCIÓN	INTERPRETACIÓN	RIESGO		
A.3.3.1	ETAPA EN LA QUE ENCUENTRA	0	ETAPA 1 O 2	ETAPAS CONTACTO - GROOMING			
		TOTAL	0				

Fuente: Elaboración propia

CAPITULO V

Marco legal

Las llamadas tecnologías de la información y la comunicación (TIC) han supuesto una verdadera revolución. Las TIC permiten un posible uso inadecuado, lo que ha llevado al legislador penal a incluir nuevos tipos que hagan frente a estos usos. Entre estos posibles usos, como ya hemos descrito en capítulos anteriores, es el llamado “*child grooming*” o ciberacoso sexual infantil, que constituye el objeto de nuestro análisis. El legislador penal ha previsto el *grooming* como figura delictiva en el artículo 183 bis del Código Penal, a través de la reforma operada por la Ley Orgánica 5/2010, integrado en la órbita de los delitos contra la libertad e indemnidad sexuales.

5.1. Las TIC y el Derecho Penal: Child Grooming y otras Formas de Abuso

Las TIC suponen la posibilidad de comunicación e información sin límites, aunque no por ello pierden su carácter de inocuidad. Resulta evidente que este contexto hace bien posible el ataque a derechos fundamentales. Por este motivo, algunos autores estudian la utilización de las nuevas tecnologías incluyéndolas en la teoría político-criminal de la “*sociedad del riesgo*”, término que refiere a la posibilidad que ofrece internet para comunicar a las personas, a la vez que permite la publicación de un gran número de datos personales que no necesariamente van a ser utilizados por terceros de una forma adecuada. Hoy por hoy, los derechos cuyo riesgo puede apreciarse con mayor facilidad son la intimidad y la propia imagen (en este sentido, no olvidemos la posibilidad de subir fotografías y vídeos en los que aparecen personas que no necesariamente han consentido la publicación de la imagen o que se mencione su nombre), por lo que es esencial que en las redes sociales se respete lo legislado en la Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal.

Recientemente han aparecido un conjunto de anglicismos que hacen referencia a diferentes fenómenos de abuso en los que están implicados menores y que podrían concretarse en: *ciberbullying*, el *happy slapping*, el *sexting* y el *grooming*. El Instituto Nacional de Tecnologías de la Comunicación (INTECO) define el **ciberbullying** como “*el uso y difusión de información lesiva o difamatoria en formato electrónico a través de medios de comunicación como el correo electrónico, la mensajería instantánea, las redes sociales, la mensajería de texto a través de teléfonos o dispositivos móviles o la publicación de vídeos y fotografías en plataformas electrónicas de difusión de contenidos*” [22]. El **happy slapping** podría traducirse como “cachete feliz”. Se trata de un fenómeno bastante nuevo e implica una agresión rápida y sin motivo alguno. Todo este proceso se graba y se publica en internet para darle difusión. A diferencia del ciberbullying, consiste en una agresión única entre desconocidos de forma sorpresiva sin componente sexual. Se trata del uso violento de la tecnología con fines lúdicos. El *sexting* es un abuso mediante TIC que incluye un elemento sexual, y que proviene de una unión entre *sex* (sexo) y *texting* (comunicación mediante mensajes de texto o SMS). Definida por Bartrina Andrés como “*difusión o publicación de contenidos (principalmente fotografías o vídeos) de tipo sexual, producidos por el mismo remitente utilizando su móvil u otro dispositivo tecnológico*” [23]. Se trata de una conducta que deriva de un primer comportamiento inicial de la persona que posteriormente se convierte en víctima. Una conducta que se inicie como sexting puede llegar a implicar otro abuso (esencialmente grooming o ciberbullying). La **pornografía infantil** se ha definido como “*toda aquella representación visual y real de un menor desarrollando actividades sexuales explícitas; esto es, en donde aparezca contacto sexual (incluyendo el genital-genital, oral-genital, anal-genital u oral-anal entre menores o un adulto y un menor), brutalidad, masturbación, desarrollo de conductas sádicas o masoquistas o exhibición lasciva de los genitales o el área púbica de un infante*” [24]. Se trata éste de un concepto de corte criminológico, cuyo reflejo jurídico se encuentra en el artículo 189 del Código Penal. Se podría definir como aquel material de contenido sexual (pornográfico) en el que aparezcan menores. Y para concluir, está el **Grooming**, término definido por INTECO como “*las acciones realizadas deliberadamente para establecer una relación y un control emocional sobre un niño o niña con el fin de preparar el terreno para el abuso sexual del menor*” [22]. Como ya conocemos de capítulos anteriores, el proceso de Grooming se ve facilitado por

las nuevas tecnologías y las TIC que además de facilitar la mentira- internet es el paradigma de ello- otorgan una sensación de falsa seguridad que facilita, por un lado, que el agresor haga el intento y, por otro, que el menor acceda a sus proposiciones. Para finalizar, hay que tener en cuenta que hay diversas formas de actuar para prevenir y evitar el *grooming*. Hay, por un lado, respuestas legislativas: en el caso de España, la inclusión del delito de *grooming* (o, mejor dicho, de un delito que se basa en el *grooming*). Pero también hay respuestas donde se han tomado medidas no legislativas. En cuanto a estas medidas no legislativas, es necesario mencionar que las propias estructuras que favorecen el *grooming* incluyen algunas medidas (aplicadas o no) contra el *grooming* (los términos de uso de Facebook o MySpace, por ejemplo). En estos casos, lo más común es que se prohíban en los términos de uso las conductas abusivas y que, en caso de detectarse, se bloquee al usuario que las lleva a cabo y se dé de baja la cuenta mediante la que se ha producido el abuso. A ello hay que añadir las iniciativas de algunas asociaciones (el botón de “*report abuse*” es una de ellas), que permiten la protección de los menores del abuso -no solamente del *grooming*, sino de todos los comportamientos que puedan considerarse abusivos-. En el caso concreto del botón “*report abuse*” se permite que el menor que se considere víctima, aunque también cualquier otra persona que acceda a contenidos que considere abusivos, denuncie el contenido abusivo de forma anónima. Con ello, aunque no se retira inmediatamente el contenido, se permite que el proveedor analice el contenido y si, efectivamente, puede considerarse abusivo para, en su caso, retirarlo del acceso público.

5.2. El delito de *Grooming* (art. 183 bis del Código Penal)

5.2.1. El artículo 183 bis

Una de las más llamativas reformas llevadas a cabo por la Ley Orgánica 5/2010 de 22 de junio, por la que se modifica la Ley Orgánica 10/1995, de 23 de noviembre, del Código Penal es la introducción *ex novo* de un artículo 183bis, con el siguiente tenor literal:

“El que a través de Internet, del teléfono o de cualquier otra tecnología de la información y la comunicación contacte con un menor de trece años y proponga concertar un encuentro con el mismo a fin de cometer cualquiera de los delitos descritos en los artículos 178 a 183 y 189, siempre que tal propuesta se acompañe de actos materiales encaminados al acercamiento, será castigado con la pena de uno a tres años de prisión o multa de doce a veinticuatro meses, sin perjuicio de las penas correspondientes a los delitos en su caso cometidos. Las penas se impondrán en su mitad superior cuando el acercamiento se obtenga mediante coacción, intimidación o engaño”.

El artículo 183 bis es, en concreto, el artículo que se encarga de criminalizar la conducta de “*child grooming*” en el Código Penal español (CP). El delito de *grooming* no se había incluido en el Código hasta el 2010 y, en concreto, hasta la reforma que en él introduce la Ley Orgánica (LO) 5/2010 (Ley Orgánica 5/2010, de 22 de junio, por la que se modifica la Ley Orgánica 10/1995, de 23 de noviembre, del Código Penal. Publicada en BOE núm. 152, de 23 de junio de 2010). Se trata por lo tanto de una incorporación reciente. No obstante, se trata de un delito

que no se preveía en la iniciativa original de la reforma del Código Penal. Dos son las razones esenciales por las que el legislador justifica la inclusión: por un lado la especial situación de vulnerabilidad en la que se encuentran los menores de trece años y el doble daño que reciben cuando son víctimas de un delito sexual. Por otro, el aumento en el uso de internet para realizar conductas sexuales contra menores y, en concreto, *grooming* en sentido amplio.

5.2.2. Requisitos típicos del delito de grooming

En el presente apartado se va a exponer la estructura del delito de *grooming* bajo la configuración que el legislador ha dado al delito en el artículo 183 bis CP. Se trata de un delito con una estructura ciertamente compleja, y que algunos autores han considerado como tipo mixto acumulativo [25] [26] mientras otros consideran esta afirmación desacertada, pues se trata de un tipo compuesto, que requiere la realización de diversas conductas sin que ninguna de ellas, por sí misma, constituya un delito (a diferencia del tipo mixto acumulativo que implicaría el ataque a diversos bienes jurídicos y, por tanto, la realización de diferentes delitos). Lo que queda claro, en cualquier caso, es que el tipo requiere la realización de diversos actos para poder ser calificado como delito de *grooming*. En primer lugar, se requiere un contacto con un menor de trece años a través de una TIC. Parte de la doctrina se inclina por entender que no es suficiente con que haya un intento de contacto, sino que además es necesario que el menor conteste a esta petición de contacto, aunque también hay autores que entienden que es suficiente con que haya conocimiento por parte del menor; concluimos afirmando que el contacto debe llevarse a cabo a través de una TIC, definida en sentido amplio, ya que el literal del Artículo 183 bis CP incluye expresamente una referencia a internet y teléfono a lo que añade “*cualquier otra tecnología de la información y de la comunicación*”.

En segundo lugar, se exige que haya una propuesta de encuentro para cometer contra el menor un delito sexual. En cuanto a la propuesta de encuentro, algunos autores han considerado que no es necesario que se trate de un encuentro “físico” sino que basta con un encuentro “virtual”. Efectivamente, actualmente el uso de las tecnologías son totalmente habituales los encuentros virtuales. Es necesario que el autor tenga la intención de cometer uno de los delitos previstos en los artículos 178 a 183 y 189 del CP. Así, se habla de un “*dolo reduplicado*” y de la exigencia de un “*ánimo tendencial*”. Se trata de una cuestión verdaderamente difícil de probar, como es la intención interna del autor y que, de hecho, raramente podrá ser objeto de auténtica prueba, sino de una prueba indiciaria, a lo sumo.

Por último, se exige que la propuesta se acompañe de actos materiales de acercamiento. Se trata de un requisito que implica que se haga patente el riesgo para el bien jurídico, de forma que la amenaza a la que se vea sometido sea más cercana. Gran parte de la doctrina entiende que se trata de un concepto jurídico indeterminado, que sólo puede concretarse caso a caso. Así, no se ha querido limitar este requisito y, simplemente, se ha explicitado su naturaleza (deben ser materiales) y su finalidad (el acercamiento).

5.3. Conclusiones

Con la exposición anterior, la conclusión directa es que, efectivamente, las TIC pueden ser una fuente de riesgo en cuanto a posibles comportamientos lesivos. Es evidente que las TIC abren las puertas a la realización de diferentes delitos por dos razones: favorecen el anonimato, por un lado, y aseguran una falsa sensación de seguridad (tanto para el agresor que es quien realiza el *grooming* como para la víctima, el menor), de forma que el efecto psicológico que se genera es el de la seguridad propia de quien está a cubierto. Este posible uso malintencionado favorecido por las TIC es el que ha servido al legislador para justificar su inclusión como nuevo delito, que se ha fundamentado como reacción frente al mal uso de las TIC que han llevado a victimizar a menores quienes, de acuerdo con el legislador, se encuentran en una posición de inferioridad y falta de protección.

Al introducir el nuevo delito de *grooming*, se pretende proteger la llamada *indemnidad sexual* de los menores de trece años, entendida ésta como la adecuada formación de la personalidad del menor en materia sexual. La tipificación del delito de *grooming* se ha configurado como la criminalización de actos preparatorios que, de otro modo serían impunes (y así lo entiende gran parte de la doctrina). Sin embargo, el legislador ha caído en el derecho penal simbólico. Es decir, ha criminalizado una conducta que, pese a que socialmente pueda percibirse como un problema, es muy cuestionable a nivel jurídico, con el objetivo de dar la imagen de la solución de un problema al que, por otra parte, ha dado una solución no del todo adecuada, siendo más necesario asegurar la educación de los menores para evitar que puedan ser víctimas. En todo caso, el marco legal donde se encuadra el “Child Grooming” es muy complejo y, de hecho, muy raramente aplicado, como se aprecia en el número de ocasiones - sólo se han encontrado varias sentencias- en las que se ha condenado por el delito de *grooming* en sus casi seis años de vigencia.

Capítulo VI

Presupuesto y planificación

En este capítulo se presenta la planificación y fases de desarrollo del presente Proyecto Fin de Carrera así como el presupuesto del mismo, desglosándolo en los correspondientes costes de material y costes de recursos humanos.

6.1 Planificación

A continuación se resumen las principales fases del proceso de desarrollo llevado a cabo en el presente proyecto, identificándose las tareas más significativas en cada una de ellas:

Fase I. Análisis del estado del arte

- Obtener una visión global de las tecnologías y herramientas relacionadas con el presente Proyecto Fin de Carrera. Varias iniciativas se han desarrollado para la protección de los niños y sus padres.
- Revisión sobre las distintas técnicas de análisis de conversaciones aplicadas en la detección de pedófilos en Internet:
 - **Técnicas en el procesado y Análisis Léxico de actos de diálogo:**
 - Arquitectura de agentes inteligentes para el procesamiento de información en la detección de explotación de niños
 - Categorías de actos de Diálogos en conversaciones de chat
 - Técnicas de categorización automática de conversaciones con uso del “corpus” disponible en www.perverted-justice.com sobre conversaciones pedófilas.
 - Otras herramientas de muestreo de conversaciones en redes sociales: Butterfly, ChatTrack y PieSpy.
 - **Agentes software para la detección de conversaciones con contenido pedófilo:**
 - Diseño de Sistemas Inteligentes para la detección de actos de diálogo con posible contenido pedófilo
 - SWEETIE
 - NEGOBOT

Fase II. Diseño y Desarrollo de la aplicación

- **Análisis introductorio del contexto de la pedofilia:**
 - Definición
 - Indicadores teóricos psicológicos de la Psicopatología del pedófilo
 - Etiología
- **Análisis de los perfiles usados en la elaboración del Modelo de Riesgos para la detección de conversaciones con contenido pedófilo:**
 - Análisis perfil **Víctima** y desarrollo para incorporación en el Modelo de Riesgos.
 - Análisis perfil **Pedófilo** y desarrollo para incorporación en el Modelo de Riesgos.
 - Análisis perfil **Conversación** y desarrollo para incorporación en el Modelo de Riesgos.

Fase III. Implementación del Modelo de Riesgos

- Operativización de las variables implicadas en el desarrollo de cada uno de los perfiles en el Modelo de Riesgos.
- Generación del Modelo de Riesgos de detección de actos de diálogo con contenido pedófilo.

Fase IV. Validación

- Diseñar un ejemplo práctico de unidad conversacional con posible contenido pedófilo.
- Realización de diferentes pruebas con diferentes entradas de la unidad conversacional tomada como referencia para cada uno de los perfiles diseñados en nuestro Modelo de Riesgos.

FASE V. Despliegue y puesta en marcha

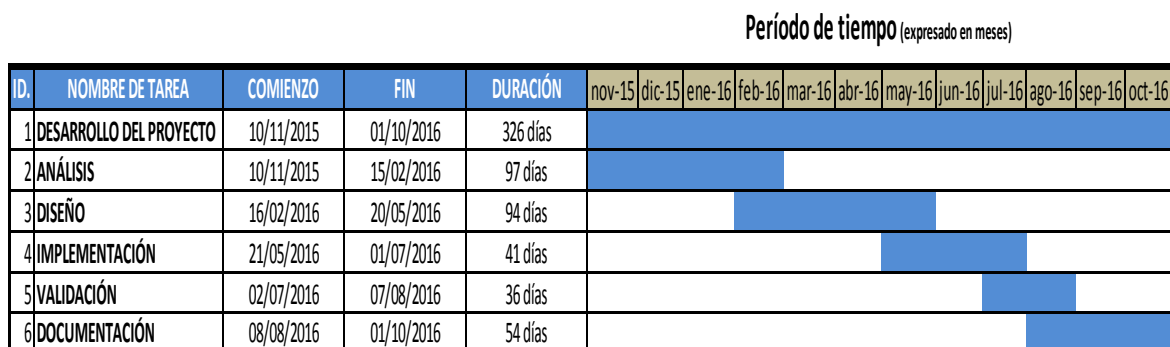
- Esta fase, que excede el objetivo de este proyecto, pasaría por la integración del Modelo de Riesgos en un sistema inteligente de detección de conversaciones con contenido pedófilo.

Fase V. Documentación

- Generar un documento que describa el trabajo realizado.
- Generar un documento que permita presentar los principales hitos de este proyecto.

La planificación de la ejecución de estas fases se muestra bajo un diagrama de Gantt en la Figura 18.

Figura 18. Diagrama de Gantt



Fuente: Elaboración propia

6.2 Costes

Una vez establecida la planificación, procedemos a realizar una estimación del presupuesto, en el que deben tenerse en cuenta diferentes tipos de costes:

COSTES DE MATERIAL: el coste de los equipos, el coste de las licencias de software y/o hardware.

COSTES DE PERSONAL: incluyen el coste del personal con cargo al proyecto, el coste de las posibles subcontrataciones, etc.

COSTES INDIRECTOS: tales como gastos en desplazamiento, dietas, luz, material de oficina, etc.

6.2.1 Costes de material

Para la realización de este Proyecto Fin de Carrera se ha requerido el uso de un ordenador portátil y de un monitor y del correspondiente software (MS Office 2010). Se considera que ambos equipos tienen una vida útil de 48 meses cuyos cálculos de amortización a 326 meses son presentados en la Tabla 2.

Tabla 2. Costes de material

Concepto	Detalle	Unidades	Coste/Unidad	Amortización (%)	Total
HARDWARE	Ordenador	1	1.000,00 €	10	100,00 €
	Monitor	1	400,00 €	10	40,00 €
SOFTWARE	Licencia MS Office	1	500,00 €	10	50,00€
					Total: 190,00 €
					Total con IVA: 229,90 €

Fuente: Elaboración propia

6.2.2 Costes de recursos humanos

Para el cálculo de costes de recursos humanos se considera la suma de horas que han sido empleadas para la implementación del proyecto. Para el caso que nos ocupa el proyecto ha sido realizado íntegramente por una Ingeniera técnica en Informática de gestión cuyo coste hora estimamos en 20 euros.

De esta forma, como se muestra en la Tabla 3, los costes incurridos en cada fase del proyecto vienen determinados por el número de horas invertidas en cada fase, atendiendo a la planificación realizada anteriormente, considerando una semana laboral de 5 días y una jornada de 8 horas/día.

Tabla 3. Desglose costes de recursos humanos del proyecto

Tarea	Horas trabajadas	Coste/Hora	Total
Análisis	40	20,00 €	
Diseño	70		
Desarrollo	60		
Implementación	45		
Validación	25		
Documentación	50		
Total	290		5.800,00 €

Fuente: Elaboración propia

6.2.3 Costes indirectos

Tabla 4. Costes indirectos asociados al proyecto

Concepto	Detalle	Unidades	Coste/Unidad	Amortización (%)	Total
Material	Material oficina	1	300,00 €	10	30,00 €
Viajes	Gasto transporte propio y público	30	10,00 €	10	30,00 €
Suministros	Luz	1	250,00 €	10	25,00 €
					Total: 85,00 €
					Total con IVA: 102,85 €

Fuente: Elaboración propia

6.2.4 Presupuesto Total del Proyecto y Ganancia Estimada

Tabla 5. Presupuesto total y ganancia estimada del proyecto

COSTES	TOTAL
COSTE MATERIAL	229,90 €
COSTE RECURSOS HUMANOS	5.800,00 €
COSTES INDIRECTOS	102,85 €
TOTAL:	6.132,75 €
% GANANCIA ESTIMADA	10%
IMPORTE GANANCIA	613,275 €

Fuente: Elaboración propia

CAPITULO VII

Conclusiones y líneas de trabajo futuras

En este capítulo se exponen las conclusiones obtenidas tras la ejecución del presente Proyecto Fin de Carrera. Además, se proponen una continuación del proyecto en cuestión en un futuro.

7.1 Conclusiones

En el presente Proyecto Fin de Carrera, partiendo de un análisis del perfil y modus operandi del pedófilo, víctima y, conversación como agente de interacción entre ambos, se ha diseñado un Modelo de Riesgos, a nivel conceptual, de detección de conversaciones con contenido pedófilo en comunidades on-line (redes sociales). El objetivo, el diseño de este módulo de detección con una probabilidad eficaz de poder detectar actos de diálogo con contenido pedófilo en las redes sociales, se ha realizado en base a un estudio profundo del perfil de nuestras variables, pedófilo, víctima y conversación; por lo tanto este proyecto parte de hipótesis contrastadas con el estudio realizado de la materia en base a los datos de los que se disponía, pero sería recomendable, e incluso sería una nueva redefinición y mejora de este módulo, ampliarlo haciendo uso de una metodología experimental de estudio de campo, con todo lo que ello conlleva, entrevistas, cuestionarios, etc., que conferiría un carácter absolutamente científico a nuestro proyecto, paso previsto para líneas de trabajo futuras.

7.2 Líneas de trabajo futuras

El presente Proyecto Fin de Carrera forma parte como primer módulo de un proyecto más ambicioso y que además de abarcar como primera fase la detección de conversaciones con contenido pedófilo en redes sociales con una alta eficacia y probabilidad, comprende un segundo módulo de implementación de un producto software cuya función sea no sólo la de proteger, a nivel control parental, sino actuar de forma que pueda ser utilizada como herramienta de control de uso de las conductas abusivas, bloqueando el proceso de acto de diálogo (bien sea por medio del bloqueo del pedófilo, o de la víctima, o ambos) y consecuentemente alertando a las autoridades competentes sobre un delito de Grooming.

A tal fin dicha implementación se encuentra en fase de desarrollo con técnicas SVMs, como se ha mencionado anteriormente.

Bibliografía

- [1] R. V. Krafft-Ebin, *Psychopathia Sexualis: The Classic Study of Deviant Sex*, New York: Arcade Publishing, 1965.
- [2] A. P. Association, *DSM IV TR. Manual Diagnóstico y estadístico de los trastornos mentales*. Texto revisado, Barcelona: Masson, 2002.
- [3] J. Murray, «Psychological profile of pedophiles and child molesters,» *Journal of Psychology*, vol. 134, nº 134, pp. 211-224, 2000.
- [4] L. J. Cohen, S. W. Gans, P. G. McGeoch, O. Poznansky, Y. Itskovich y S. Murphy, «Impulsive personality traits in male pedophiles versus healthy controls: is pedophilia a impulsive-aggressive disorder?,» *Comprehensive Psychiatry*, vol. 43, nº 43, pp. 127-134, 2002.
- [5] R. Eisenman, «Denigration of a victim in group psychotherapy by violent vs seductive sex offenders,» *Psychological Report*, vol. 72, nº 72, pp. 413-414, 1993.
- [6] R. Dickey, D. Nussbaum, K. Chevolleau y H. Davidson, «Age as a differential characteristic of rapists, pedophiles, and sexual sadists,» *Journal of Sex and Marital Therapy*, vol. 28, nº 28, pp. 211-218, 2002.
- [7] I. B. A Moller, «Sexual abuse of children- general practice of forensic perpetrator

- assessment,» *Gesundheitswesen*, vol. 56, nº 56, pp. 47-50, 1994.
- [8] M. R. Bridges, J. S. Wilson y C. B. Gacono, «A Rorschach investigation of defensiveness, self-perception, interpersonal relations, and affective states in incarcerated pedophiles,» *Journal of Personality Assessment*, vol. 70, nº 70, pp. 365-385, 1998.
- [9] B. Glaser, «Psychiatry and paedophilia: a major health issue,» *Australian and New Zealand Journal of Psychiatry*, vol. 32, nº 32, pp. 162-167, 1998.
- [10] S. S. Johnston, A. P. French, W. F. Schouweiler y F. A. Johnston, «Naiveté and need for affection among pedophiles,» *Journal of Clinical Psychology*, vol. 48, nº 48, pp. 620-627, 1992.
- [11] K. F. Durkin, «Misuse of the Internet by Pedophiles: Implications for Law Enforcement and Probation Practice,» *Federal Probation*, vol. September, nº 61, pp. 14-18, 1997.
- [12] A. Ankou, G. Romero y G. Mauri, «Design of the overall system architecture, Tech. Rep. 3.1,» de *OPEN meter Consortium*, 02 02 2010.
- [13] C. H. Martell y E. N. Forsythand, «Lexical and discourse analysis of online chat dialogo,» *Proc. International Conference on Semantic Computing ICSC*, pp. 19-26, 2007.
- [14] J. Lin, *Automatic author profiling of online chat logs*, N. P. School, Ed., 2007.
- [15] N. Pendar, *Toward spotting the pedophile telling victim from predator in text chats*, Proc. International Conference on Semantic Computing ICSC, 2007, pp. 235-241.
- [16] H. L. Neil, W. Van Dyke y P. Maes, *Butterfly: a conversation-finding agent for Internet relay chat Irvy*.
- [17] J. Bengel y et al, «Chattrack: chat room topic detection using clasification,» de *In 2nd Symposium on Intelligence and Security Informatics*, 2004.
- [18] P. Mutton, «www.jibble.e.org/piespy,» 2004. [En línea].
- [19] C. Laorden Gómez, P. Galán García, I. Santos Grueiro, P. Garcia Bringas y B. Sanz Urquijo, «Negobot: Agente conversacional basado en teoría de juegos para la detección de conductas pedófilas,» *DeustoTech Computing*, pp. 64-67, 2016.
- [20] E. Echeburúa y C. Guerricaechevarría, *Abuso sexual en la infancia: víctimas y agresores*, Barcelona: Ariel, 2000.
- [21] R. Braschler, «Camphow.com,» 2014. [En línea]. Available: <http://www.camphow.com/camping/risk-management/stages-child-grooming-abuse-prevention-detection-response/trackback>. [Último acceso: marzo 2016].
- [22] I. Observatorio de la seguridad de la información. Área Jurídica dela Seguridad y las TIC. Instituto Nacional de Tecnologías de la Comunicación, *Guía Legal sobre Cyberbullying y Grooming*.
- [23] M. J. Bartrina Andrés, *Anàlisi i abordatge de l'assetjament entre iguals mitjançant l'ús*

de les noves tecnologies. Justícia juvenil i adolescents en l'era digital, 2011.

- [24] D. L. Morillas Fernández , Análisis dogmático y criminológico de los delitos de pornografía infantil, Madrid: Dykinson, 2006.
- [25] M. M. González Tascón, «El nuevo delito de acceso a niños con fines sexuales a través de las TIC,» *Estudios Penales y Criminológicos*, vol. XXXI , p. 245, 2011.
- [26] F. Perez Ferrer, «El nuevo delito de ciberacoso o child grooming en el Código Penal español (artículo 183 bis),» *Diario La Ley, sección doctrina Nº 7915*, p. 5, 2012.
- [27] C. Schinaia, Pedofilia Pedofilias "El psicoanálisis y el mundo del pedófilo", APM, 2011.
- [28] W. Marshall, Agresores sexuales, Barcelona: Ariel, 2001.
- [29] B. Glaser, «Psychiatry and paedophilia: a major health issue,» *Australian and New Zealand Journal of Psychiatry*, vol. 32, nº 32, pp. 162-167, 1998.
- [30] K. Freund y M. Kuban , «Toward a testable developmental model of pedophilia: the development of erotic age preference,» *Child Abuse and Neglect*, vol. 17, pp. 315-324, 1993.
- [31] J. R. Feierman, Pedophilia Bisocial Dimensions, Springer-Verlag.
- [32] L. J. Cohen, P. G. McGeoch, S. W. Gans, K. Nikiforov, K. Cullen y I. I. Galynker, «Childhood sexual history of 20 male pedophiles vs. 24 male healthy control subjects,» *Journal of Nervous and Mental Disease*, vol. 190, nº 190, pp. 757-766, 2002.
- [33] A. Beltran y S. Ordoñez, «Sistema inteligente para la detección de diálogos con posibles contenidos pedofílicos,» *Revista Virtual Universidad Católica del Norte*, vol. 42, pp. 164-181, 2014.
- [34] A. Beltran, «<https://www.researchgate.net/publication/259799455>,» Junio 2010. [En línea].
- [35] R. Baker, K. J. Wininger y F. Ellinston, Philosophy and Sex, Prometheus Books.